**Data Manipulation and Cleaning work**

**In Excel:**

**ACCOUNT Table:**

1. Convert the Date attribute into a yyyy-mm-dd by adding **24** in year format in Excel or SQL

=DATE(LEFT(D2, 2) + 24, MID(D2, 3, 2), RIGHT(D2, 2))

* 1993 -> 2017
* 1994 -> 2018
* 1995 -> 2019
* 1996 -> 2020
* 1997 -> 2021

1. Replace in frequency attribute “POPLATEK MESICNE” AS **Monthly Issuance**, “POPLATEKTYDNE” AS **Weekly Issuance,** and “POPLATEK POBRATU” AS I**ssuance After a Transaction** in Excel or create a case statement in SQL.
2. 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 **Sliver**, “Classic” as **Gold**,

And “Gold” as **Diamond** by using replace in Excel or by using update in SQL.

1. Convert issued attribute into yyyy-mm-dd adding 24 in year.

**CLIENT Table:**

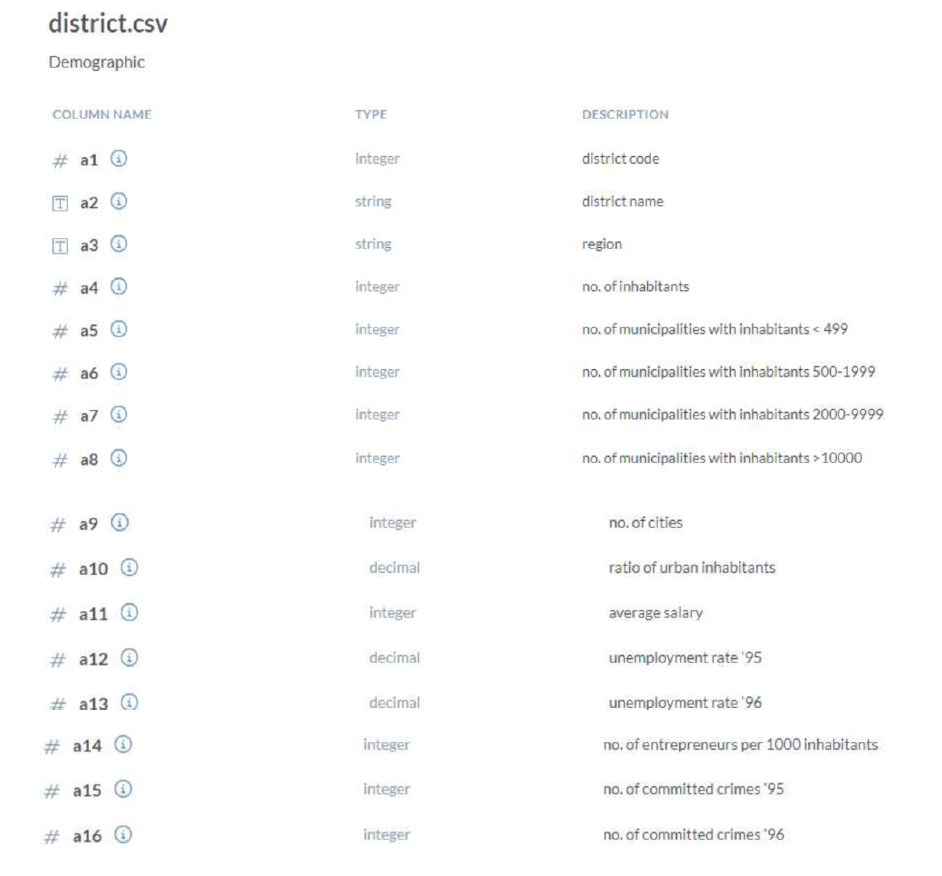
1. Convert the bith\_number attribute to yyyy-mm-dd format and also create another column named sex by applying in bith\_number 0 for females and 1 for males.

**(=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.**

**DISTRICT Table:**

1. Change all column names as shown below and delete the attributes **a12.**



**LOAN Table:**

1. Convert the Date Attribute into yyyy-mm-dd format adding 24 in year.
2. Convert “Status” Attribute value as shown below

* “A” -> Contract Finished
* “B” -> Loan Not Paid
* “C” -> Running Contract
* “D” -> Client in debt.

**In Snowflake (SQL):**

**TRANSACTIONS Table:**

1. The client wants to update the Date information in the Transaction table as follows.
2. In the Transactions table whosoever transaction count is highest, sort it in descending order and change the year from 2024,2023,2022... as follows.

* CONVERT 2021 TXN\_YEAR TO 2024
* CONVERT 2020 TXN\_YEAR TO 2023
* CONVERT 2019 TXN\_YEAR TO 2022
* CONVERT 2018 TXN\_YEAR TO 2021
* CONVERT 2017 TXN\_YEAR TO 2019
* CONVERT 2016 TXN\_YEAR TO 2020

**e.g:** UPDATE TRANSACTIONS

SET DATE = DATEADD(YEAR, 3, DATE)

WHERE YEAR(DATE) = 2021;

1. Client wants to update the NULL values for the attribute 'BANK' in the TRANSACTION table for the particular transaction year as follows.
2. In the Transactions table whosoever transaction count is highest (in the context of NULL values in the attribute BANK), sort it in descending order and change the NULL values in attribute 'BANK' for the particular transaction year in the respective order –

* 'Sky Bank', 'DBS Bank', 'Northern Bank', 'Southern Bank', 'ADB Bank'.

1. From the result, as per requirement we need to change the null values as follows.

* 2024 (226491) --> Sky Bank
* 2023 (210877) --> DBS Bank
* 2022 (0) --> No NULL values
* 2021 (99074) --> Northern Bank
* 2019 (68763) --> Southern Bank
* 2020 (22512) --> ADB Bank

**e.g:** UPDATE TRANSACTIONS

SET BANK = 'Sky Bank'

WHERE BANK IS NULL AND YEAR(DATE) = 2024;

**CARD Table:**

1. The client wants to update the Issued Date information in the Card table as follows.
2. In the CARD table whosoever issued card count is highest, sort it in descending order and change the year from 2024,2023,2022... as follows.

* CONVERT 2021 TXN\_YEAR TO 2024
* CONVERT 2020 TXN\_YEAR TO 2023
* CONVERT 2019 TXN\_YEAR TO 2022
* CONVERT 2018 TXN\_YEAR TO 2021
* CONVERT 2017 TXN\_YEAR TO 2019
* CONVERT 2016 TXN\_YEAR TO 2020

**ACCOUNT Table:**

1. The client wants to update the DATE attribute in the ACCOUNT table by adding two years to the current date.

**e.g:** UPDATE ACCOUNT

SET DATE = DATEADD(YEAR, 2, DATE)

WHERE YEAR(DATE) = 2020;

**LOAN Table:**

1. The client wants to update the DATE attribute in the LOAN table as follows.

* CONVERT 2021 YEAR TO 2024
* CONVERT 2020 YEAR TO 2023
* CONVERT 2019 YEAR TO 2022
* CONVERT 2018 YEAR TO 2021
* CONVERT 2017 YEAR TO 2019
* CONVERT 2016 YEAR TO 2020

**CLIENT Table:**

1. The client wants to add the “**AGE**” column to the CLIENT table by subtracting the birth year and last transaction year.

**e.g:** ALTER TABLE CLIENT

ADD COLUMN AGE INT;

UPDATE CLIENT

SET AGE = DATEDIFF(YEAR, BIRTH\_DATE, ‘Last Txn Date’);