

## Part 1: Data extraction and transformation

**Scenario:** You have two source systems:

1. An Oracle database containing transactions and accounts data. Tables:

- Transactions

**transaction\_number:** *Unique identifier.*

**transaction\_date:** *Date the transaction took place*

**debit\_credit:** *An indicator regarding the symbol (+/-) of the transaction. If the flag is equal to C then the amount is positive (+) and this means that money is deposited to the account, if its D then the amount is negative (-) and means that the amount is being withdrawn from the account.*

**transaction\_amount:** *The amount of the transaction.*

**transaction\_account\_number:** *The customer account who performed the transaction.*

- Accounts

**account\_number:** *The customer account number.*

**customer\_group\_id:** *This field is called CIF, and it's a group id attached to an account that can be assigned to multiple customer profiles.*

**product\_code:** *The type of the account.*

**account\_designation:** *The name on the account.*

- Profile\_groups

**customer\_group\_id:** *This field is called CIF, and it's a group id attached to an account that can be assigned to multiple customer profiles.*

**profile\_number:** *The profile id of the customer (its unique).*

2. An SQL Server database containing customer data.

- Banking\_customers

**customer\_pk:** *A unique identifier of the customer on SQL Server.*

**customer\_profile:** *The profile id of the customer (its unique).*

**salutation:** *The salutation of the customer.*

- Physical\_customers

**customer\_pk:** *A unique identifier of the customer on SQL Server.*

**customer\_firstname:** *Customers first name.*

**customer\_lastname:** *Customers last name.*

- Corporate\_customers

**customer\_pk:** *A unique identifier of the customer on SQL Server.*

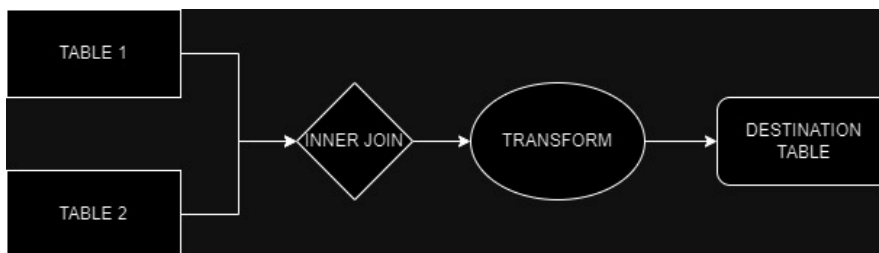
**company\_title:** *Company name.*

### Requirements:

- Extract transaction data from the Oracle database and customer data from the SQL Server database.
- Load the transformed data into a data warehouse in the SQL Server database. Create a data mart regarding transactional data for each type of customer (corporate and physical).

### Deliverables:

1. A data flow diagram of an ETL processing the above requirements. You can use <https://app.diagrams.net> to create the diagram.



2. A brief document explaining the ETL process and any challenges faced.
3. Explain the structure of the data mart and explain the pros and cons relative to alternative structures.

## Part 2: BI Engineering

### Scenario:

1. Create a report that shows the balances of each account per month assuming that the accounts were all opened on the 1<sup>st</sup> January 2023 (0 balance at that date).
2. Create a report showing all transactions of a customer for a specific date range.
3. Find the month with the most deposits sum for the entire year.

### Requirements:

- Develop SQL Server queries for the above.
- The reports should include parameters to filter by date range, month, customer account and customer name depending on the scenario.

### Deliverables:

1. Three SQL Server queries.
2. A brief document explaining the report design and how to use the parameters.
3. Explain how the previously created data mart helped you into creating the three reports.