Data Dictionary

???

Table Name	Column Name
Transactions Fact Table	TransactionID
Transactions Fact Table	TransactionStartDatetime
Transactions Fact Table	TransactionEndDatetime
Transactions Fact Table	Cardholder ID
Transactions Fact Table	Location ID
Transactions Fact Table	Transaction Type ID
Transactions Fact Table	Transaction Amount
Location Dimension Table	Location ID
~~Location Dimension Table~~	~~Location Name~~
~~Location Dimension Table~~	~~No of ATMs~~ ???
Location Dimension Table	City
Location Dimension Table	State
Location Dimension Table	Country
Customers Dimension Table	Cardholder ID
~~Customers Dimension Table~~	~~First Name~~
~~Customers Dimension Table~~	~~Last Name~~
~~Customers Dimension Table~~	~~Gender~~
~~Customers Dimension Table~~	~~ATM ID~~
~~Customers Dimension Table~~	~~Age~~
~~Customers Dimension Table~~	~~Occupation~~
Customers Dimension Table	Account Type
~~Customers Dimension Table~~	~~Is Wisabi~~ ???
Transaction Type Dimension Table	TransactionTypeID
~~Transaction Type Dimension Table~~	~~Transaction Type~~
~~Hour Dimension Table~~	~~Hours~~ ???
Hour Dimension Table	Hour Start Time
Hour Dimension Table	Hour End Time
Calendar Dimension Table	Date
Calendar Dimension Table	Quarter
Calendar Dimension Table	Month
Calendar Dimension Table	Month Name
Calendar Dimension Table	Day

Yo creo que no hace falta empezar de cero para generar el PG-BD que necesitamos. Ya tenemos un punto de partida con una tamaño bueno para empezar. Mi idea es que se pueden eliminar los datos locales (strings) y dejar solo los datos tipo ID. Para el caso de los customers, construye una biyección (nombre, apellido) --> id usando una tabla de hash. Cada vez que tu script procese un par (nombre, apellido) aplica el hash y si ya estaba definido, usas el id asociado al para. Si no, defines la nueva entrada en la tabla de hash. Idem para las ciudades, generas las GPS-location, etc. Creo que es más fácil hacer esta traducción que partir de cero. Reducimos el problema a solo crear nuevas transacciones como las que comentaste que faltaban.

Calendar Dimension Table	Is Holiday
Calendar Dimension Table	Day Name
Calendar Dimension Table	Week of Year
Calendar Dimension Table	Year
Calendar Dimension Table	Start of Month

Description
Unique identifier for each transaction in the database
Datetime when the transaction started
Datetime when the transaction was completed
Unique identifier for the cardholder performing the transaction
Unique identifier for the location of the ATM where the transaction occurred
Unique identifier for the type of transaction that was performed (e.g., withdrawal, savings, balance enquiry, tra
Amount of money involved in the transaction
Unique identifier for the ATM location
Name of the bank branch where the ATM is located
Number of ATMs
City in which the ATM is located
State in which the ATM is located
Country in which the ATM is located
Unique identifier for the cardholder
First name of the cardholder
Last name of the cardholder
Gender of the cardholder (e.g., male, female, other)
Unique identifier for the ATM that the cardholder uses
Age of the cardholder
Occupation of the cardholder
Type of account that the cardholder has (e.g., savings, checking, etc.)
Boolean flag that indicates whether the cardholder is a customer of Wisabi Bank or another bank
Unique identifier for the transaction type (e.g., 1 for withdrawal, 2 for savings, 3 for balance enquiry, 4 for trans
Name of the transaction type (e.g., "withdrawal", "savings", "balance enquiry", "transfer")
Hour of the day (0-23)
Time at which the hour begins (e.g., 12:00 AM for hour 0)
Time at which the hour ends (e.g., 1:00 AM for hour 0)
Date in YYYY-MM-DD format
Quarter of the year in which the date falls (e.g., Q1 for January-March, Q2 for April-June, etc.)
Month in which the date falls (e.g., 1 for January, 2 for February, etc.)
Name of Month (E.g January, February, etc.)
Day of the week in which the date falls (e.g., Monday, Tuesday, etc.)

Boolean flag that indicates whether the date is a public holiday
Name of Day (E.g Monday, Tuesday, etc.)
Week of Year (From 1 to 54)
Year (2022)
Start of Month for each date value