

## Sample Project for Data Migration Engineers

### Introduction

The business case is very simplified in the sample database: We are administrating a pension fund, where employers can sign up their employees, which then pay into the pension plan (“making contributions”) and are able to receive pension benefits later on. Each member has an account they accrue their contributions in.

To import contribution data into our system, we receive a Microsoft Excel file exported from the client’s legacy system and transform the data into our database structure. Each client has their own way of storing data (thus, Excel exports vary per client) and each client’s data has to be transformed into our data structure.

In this scenario, the Microsoft Excel export has already been analyzed and been transposed into database tables. With a real import, these tables would be temporary.

### Setup

With this document you have received an SQL script file containing setup instructions for both the “temporary” tables and the actual tables the data needs to be inserted to. In order to set up the database, running the script should be sufficient. If the first command fails, it is because it is for the psql command line tool. Just remove it if you are executing the script in any other environment.

### The Database Structure

The table structure consists of User, Plan, Country, Province, Member, Account, Contribution and Earning. Those are a small part of our actual database (slightly cut down), which the data goes into.

Users are application users, which are needed for the CreatedBy and UpdatedBy columns in each table, which must be filled, to see who updated the record last. The system user should be used for this, since it is an import outside the application data maintenance.

Plans are the highest entity in terms of grouping dependent entities together. Members, Accounts, Contributions and Earnings all belong to plans, as the application is set up for multiple tenants. Make sure to create a plan (with any code of your choice) for this import, as all the imported data must be grouped under this plan afterwards.

Members are the employees previously mentioned that contribute to the plan. Accounts are the group across multiple Contribution and Earning records.

Earnings describe how much a person actually earned, while Contributions describe how much a person has paid into the plan. When a member pays into the plan, they are credited with service, which is used to calculate their pension.

In addition to the actual database tables, we have temporary import tables: *contributions*, *earnings*, *memberdata* and *service*.

The table *memberdata* contains data about the member and their account. Ignore anything relating to statuses or spouses. All dates are currently in the Excel format, make sure to convert these to proper PostgreSQL dates. Consider any *Unknown* date as null. The names of each member are in a “lastName, firstName” format.

The table *contributions* has the foreign key towards the member (MemberID), as well as contributions for each year. Ignore the *cwi* columns and insert the *cont* columns as contributions. Each column represents the end of the year, so the 2010 cont would be 2010-12-31 for example.

The table *earnings* behaves basically the same, as well as the service table. Both earnings and service go into the actual Earning database table.

### Project Task

- Your task is to create an SQL script file that transforms the data in the temporary tables and inserts it into the actual tables correctly. Also include the statement(s) to create the plan.
- Mind the hints given above.
- You are free to use just plain SQL or PL/pgSQL
- There is no restriction on the database client. We recommend using psql or JetBrains's Datagrip.
- Use PostgreSQL in major version 10
- If you have any questions, feel free to reach us at [hello@g9h.co](mailto:hello@g9h.co)