Data Analysis Prompt:

1. **Data Profiling:**

The attached dataset contains duplicates, missing values, and other inconsistencies.

Run data profiling methods to identify the issues.

2. **Data Cleaning and Preprocessing:**

Identify ways to clean and prepare the data, so that the resulting dataset can be imported in a database table used for reporting. Please explain the methods used and provide a clean dataset.

3. **Data Analysis and Visualization**

Use the cleaned dataset to create a chart that displays the average client income trend over time. Expand by running demographics based on age and location.

Data Modelling

4. **Design a datamart model**
Based on the entity relationship diagram below, design an Employee star schema (1 Fact, 3 dimensions). Ensure to include measures and attributes.

5. **Data Transformation**

Based on the above transactional model, write the SQL queries required to populate the fact and dimensions designed at previous step

Reporting

6. **Data Aggregation and Reporting**

Based on the above transactional model

   a. write an SQL query that returns the total expense amounts by expense unit, expense type, and year of submission (EXPENSE_DATE)
   b. write an SQL query that returns the employees that spent more that $1000 in a year.
   c. write an SQL query that identifies all employees that spent less than $100 (or nothing at all) in 2023.