EASYPAY

Software Developer Task: Data Migration Case Study

An organization provides on-site IT technical support to individuals at their homes. The organization has been using a shared Excel sheet each year to record all work orders performed by technicians. Technicians enter data in Excel with the following fields: Technician, Notes, Total. In the Notes field, technicians fill in all details in free-text form, for example:

Technician	Notes	Total
Alban Mone	U be sherbimi i printerit dhe u ndryshua toner te klienti Andi Mucobega me daten 10/12/2024.	12000

The management has decided to digitalize the process and purchased a new ERP system. In the new system, data will be stored in a structured way across three tables: Clients, Work Orders, and Technicians. The organization has requested that all historical data inserted in Excel files be migrated to the new system.

As a developer working for the ERP provider, you need to migrate the data from Excel files into the new system.

An Excel file containing the complete client list has been provided by the Finance Department, based on data from the accounting system. The Operations Department has submitted a separate file with the work orders.

Your tasks are to:

- Import client data from the Finance file into the Clients table (ClientId, First name, Last name). Choose appropriate data types as you see fit.
- Import technicians' data from the excel into the Technicians table (TechnicianId, First name, Last name). Choose appropriate data types as you see fit.
- Import work order data such as:

Technician	Notes	Total
Alban Mone	U be sherbimi i printerit dhe u ndryshua toner te klienti Andi Mucobega me daten 10/12/2024.	12000

  into the new system. You have to extract data from the Notes column: the data for the Client name, and the date of the work order. Ensure that the extracted data is accurately mapped to the corresponding fields in the Work Orders table, by having the correct Technician and Client. You have to identify the client from the text in the Notes column.

  The new table for WorkOrders in the new system has the following fields: WorkOrderId, TechnicianId, ClientId, Information, Date, Total. Choose appropriate data types as you see fit. The Information field should contain the text as per the Notes column. The organization

has requested that you remove the client's name and date from the new Information field, as these details will be stored separately in the system and will be redundant. For example, instead of: *"U be sherbimi i printerit dhe u ndryshua toner te klienti Andi Mucobega me daten 10/12/2024.",* the preferred text for the first record's Information field would be: *"U be sherbimi i printerit dhe u ndryshua toner te klienti."*

However, this request is considered a **nice-to-have** rather than mandatory, so you may choose to skip this step if you prefer.

- Generate a CSV report summarizing the import process, clearly listing both successful and failed records.

You must consider that technicians may have made small spelling mistakes in client names. For example, the correct name according to Finance is **Andi Muçobegaj**, but the technician wrote **Andi Mucobega** in the Notes. You must account for such errors when checking names in the excel file with the correct names from the finance file.

For testing, you will be provided with synthetic data:

- An Excel file containing 106 clients.
- An Excel file with 104 sample work orders.

Keep in mind that the organization's historical data includes about 10 million work order records (stored in separate Excel files per year). For the task you should focus on a single excel work order file, but the import process must be designed for high performance to process this volume of data as quickly as possible. The organization acknowledges that the import may not be 100% flawless, but it expects a very high level of accuracy in the imported records.

Guidelines

The focus of this task is to assess both your analytical thinking and coding skills.

You must use .NET for the implementation and SQL Server for the database.

You are free to use third-party libraries, integrations or tools, as long as they are reliable and properly integrated. There are no restrictions on their usage.

While a microservice-based architecture is preferred, you may choose a different architecture if it's well-structured and justified.

The solution will be evaluated based on both technical implementation and architectural decisions.

The solution should be submitted as a Git repository. If the repository is private, please grant access to the email address from which you received this task.