

以下是一個針對初學者的「SQL Server Integration Services (SSIS) ETL」專案設計，基於提供的學習課綱，涵蓋核心概念並設計一個簡單實用的任務，讓初學者能夠逐步掌握 SSIS 的基本功能。專案旨在模擬真實場景，從資料整合到部署，提供完整的學習體驗。

專案名稱

銷售資料整合與報表準備

一、任務說明

背景

你是一家 3C 零售公司的資料分析助理，負責將分散的銷售資料整合到中央資料庫，以供後續報表分析。資料來源包括一個 SQL Server 資料表與一個 CSV 檔案，你需要使用 SSIS 設計一個 ETL 流程，提取資料、進行簡單轉換並載入目標表。

任務目標

- 學習 SSIS 的基本工具與環境。
- 設計並執行一個簡單的 ETL 封裝。
- 熟悉資料流程、運算式與錯誤處理。
- 將結果部署到 SSIS 伺服器。

輸入來源

1. SQL Server 資料表 (SalesOrders)

- 欄位：ID, CName, price, Order
- 範例資料：

| ID | CName | price | Order |
|----|-------|-------|------------|
| 1 | Alice | 500 | 2025-01-01 |
| 2 | Bob | 300 | 2025-01-02 |

2. CSV 檔案 (ExternalSales.csv)

- 欄位：SID, Customer, SPrice, Date
- 範例資料：

| SID | Customer | SPrice | Date |
|------|----------|--------|------------|
| E001 | Alice | 200 | 2025-01-03 |
| E002 | Charlie | 400 | 2025-01-04 |

預期輸出

- 目標資料表 (IntegratedSales)
 - 欄位：RecordID, CustomerName, TotalAmount, SaleDate
 - 範例結果：

| ID | CName | TotalPrice | SaleDate |
|----|---------|------------|------------|
| 1 | Alice | 700 | 2025-01-01 |
| 2 | Bob | 300 | 2025-01-02 |
| 3 | Charlie | 400 | 2025-01-04 |

- 錯誤日誌檔案 (ErrorLog.csv)
 - 儲存不符合條件的記錄 (例如金額小於 0)。

任務要求

- 使用 SSIS 提取兩個來源的資料。
- 轉換資料：統一欄位名稱，過濾無效金額 (< 0)。
- 載入目標表並記錄錯誤。
- 部署封裝並驗證結果。

二、任務執行

步驟指引

1. 環境準備
- 安裝 SQL Server (含 SSIS) 與 SQL Server Data Tools (SSDT)。
 - 在 SQL Server 中建立資料庫與表格：

```
CREATE DATABASE DB;
GO
USE DB;

CREATE TABLE TblName (

);

INSERT INTO TblName VALUES
(1, 'Alice', 500, '2025-01-01'),
(2, 'Bob', 300, '2025-01-02');

CREATE TABLE IntegratedSales (

);
```

- 準備 CSV 檔案 (ExternalSales.csv) 並儲存到本地路徑 (如 C:\Data\)。
2. 建立 SSIS 專案

- 開啟 SSDT，選擇「Integration Services Project」，命名為 `SalesETLProject`。
- 新增封裝，命名為 `SalesIntegration.dtsx`。

3. 設計 ETL 流程

- 連線管理器：
- 資料流程任務：
- 日誌記錄：

4. 測試與部署

- 在 SSDT 中按 F5 執行，檢查資料流程是否正確。
- 部署封裝：
 - 右鍵專案 > 「部署」，設定 SSIS Catalog (`localhost`) 與資料夾 (`SalesETL`)。

5. 驗證結果

- 使用 SSMS 查詢 `IntegratedSales` 是否包含正確資料。
- 檢查 `ErrorLog.csv` 是否記錄無效資料 (若有)。

三、提交文件說明

提交以下文件：

1. SSIS 封裝檔案

- 檔案名稱：`SalesIntegration.dtsx`
- 包含完整的 ETL 流程設計。

2. 目標資料表查詢結果

- 檔案名稱：`IntegratedSales_Result.txt`
- 內容：執行以下查詢並儲存結果：

```
SELECT * FROM IntegratedSales;
```

3. 錯誤日誌檔案

- 檔案名稱：`ErrorLog.csv`
- 若無錯誤，提交空檔案並註明。

4. 說明文件

- 檔案名稱：`readme.txt`
- 內容：
 - 任務簡述 (例如「此封裝整合 SQL 與 CSV 銷售資料並載入目標表」)。
 - 執行步驟 (例如「將 CSV 放在 C:\Data\，然後在 SSDT 執行」)。
 - 遇到的問題與解決方法 (若有)。

四、評量標準

總分 100 分，分為以下項目：

| 評量項目 | 說明 | 分數 |
|---------|---|------|
| 功能正確性 | 封裝能正確提取、轉換並載入資料，結果符合預期。 | 40 分 |
| 設計完整性 | 使用資料流程、運算式與錯誤處理，結構清晰並符合課綱要求。 | 20 分 |
| 部署與執行 | 封裝成功部署到 SSIS Catalog 並可執行，包含日誌記錄。 | 20 分 |
| 提交文件完整性 | 提交所有要求檔案，內容正確且格式符合規範。 | 10 分 |
| 說明文件清晰度 | <code>readme.txt</code> 清楚描述任務與步驟，易於理解。 | 10 分 |