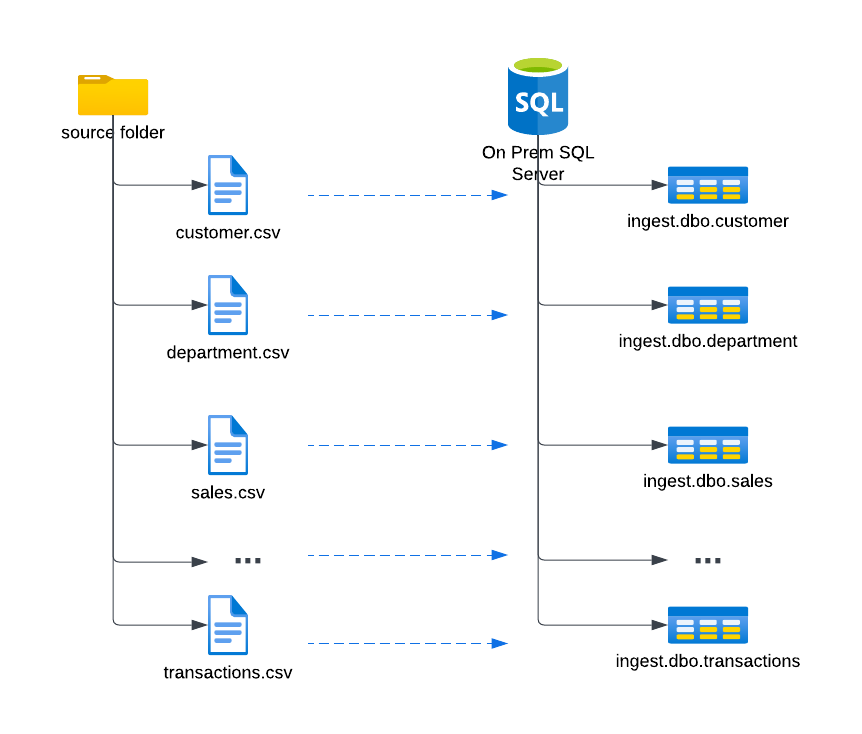
Task for Data Engineer: Designing and Implementing an ETL Process

**Objective:**  
The goal is to establish a robust Extract, Transform, and Load (ETL) process for handling daily-generated CSV files from the Point of Sale (POS) system. This task requires the development of a data integration solution that efficiently ingests data into an ingest database while ensuring security, adaptability to changing source files, and effective error handling. Additionally, a deployment and scheduling strategy must be devised for seamless execution.



**Requirements:**

1. **Dynamic Source File Handling:**
   * Develop a data integration process capable of dynamically adapting to different CSV files generated daily by the POS system.
   * Implement a mechanism to identify and handle new files with varying names, creating corresponding tables in the ingest database as needed.
2. **Security Measures:**
   * Avoid storing sensitive credentials (e.g., database login/password) directly within scripts for security reasons.
3. **Deployment Strategy:**
   * Consider leveraging deployment tools or version control systems for efficient and reproducible deployments.
4. **Scheduling:**
   * Implement a scheduling mechanism to automate the daily execution of the ETL process.
   * Utilize built-in scheduling features of ETL tools, e.g., SQL Agent, Windows Task Scheduler, Airflow and so on.
5. **Monitoring and Logging:**
   * Establish robust monitoring and logging mechanisms to track key performance metrics and identify potential issues.
   * Log relevant information for auditing purposes and facilitate troubleshooting in case of failures.
6. **Error Handling:**
   * Implement a comprehensive error-handling strategy to address potential issues during the ETL process.
   * Notify the Database Administrator (DBA) promptly in case of failures, using mechanisms such as email alerts or detailed logging.
7. **Scalability:**
   * Design the ETL solution with scalability in mind, allowing for future growth in data volumes.
8. **Documentation:**
   * Provide comprehensive documentation for deployment and scheduling procedures.
   * Document dependencies, configurations, and considerations to facilitate knowledge transfer and future maintenance.

**Tools and Technologies:**

* Developers are free to choose suitable tools and programming languages for ETL implementation, such as ADF, SSIS, Python, or a combination thereof.

**Assessment Criteria:**

* The solution should demonstrate adaptability to changing source files, security best practices, and effective error handling.
* The deployment and scheduling strategy should be well-documented, ensuring repeatability and scalability.
* Consideration of scalability, monitoring, and logging aspects will be evaluated for robustness.
* Clarity and completeness of documentation will be crucial for knowledge transfer and future maintenance.

**Submission:**

* Provide the ETL solution codebase, deployment scripts, and documentation. Include clear instructions for dependencies, configurations and any specific considerations for deploying and scheduling the ETL process.