

Übungsaufgabe: Mehrere Schritte in einem DAG

User Story 5: Laden der Daten in eine PostgreSQL-Datenbank

Beschreibung: Als Entwickler möchte ich die transformierten Daten in eine PostgreSQL-Datenbank laden, damit sie für Analysen und Berichte verfügbar sind.

Akzeptanzkriterien:

- Die transformierten Daten werden korrekt in die PostgreSQL-Datenbank geschrieben.

Schritte zur Umsetzung:

1. Eine neue Verbindung zur PostgreSQL-Datenbank erstellen:

- Erstellen Sie eine neue Verbindung zur PostgreSQL-Datenbank in der Airflow-UI.
- Die Verbindungsdaten finden Sie in der `docker-compose.yml` Datei.
- <https://www.astronomer.io/docs/learn/connections/postgres/>

2. Erstellen Sie den DAG

- Editieren Sie die Datei `minio_to_postgres_dag_exercise.py` und ergänzen Sie die fehlenden Teile.
 - Kopieren Sie die Datei in den `dags` Ordner und aktivieren Sie den DAG in der Airflow-UI.
4. **DAG ausführen:** Aktivieren Sie den DAG und führen Sie ihn aus. Überprüfen Sie die Logs im Web-Interface, ob die Datei erfolgreich transformiert wurde.

Ressourcen:

Airflow

- <https://airflow.apache.org/docs/apache-airflow/stable/core-concepts/connections.html>
- <https://www.astronomer.io/docs/learn/what-is-a-hook>
- https://airflow.apache.org/docs/apache-airflow-providers-postgres/stable/_api/airflow/providers/postgres/hooks/postgres/index.html
- <https://airflow.apache.org/docs/apache-airflow-providers-common-sql/stable/operators.html>

Python

- https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.to_sql.html
- <https://www.sqlalchemy.org>