

Lab 1 – External Table 을 이용한 테이블 생성(Polybase)

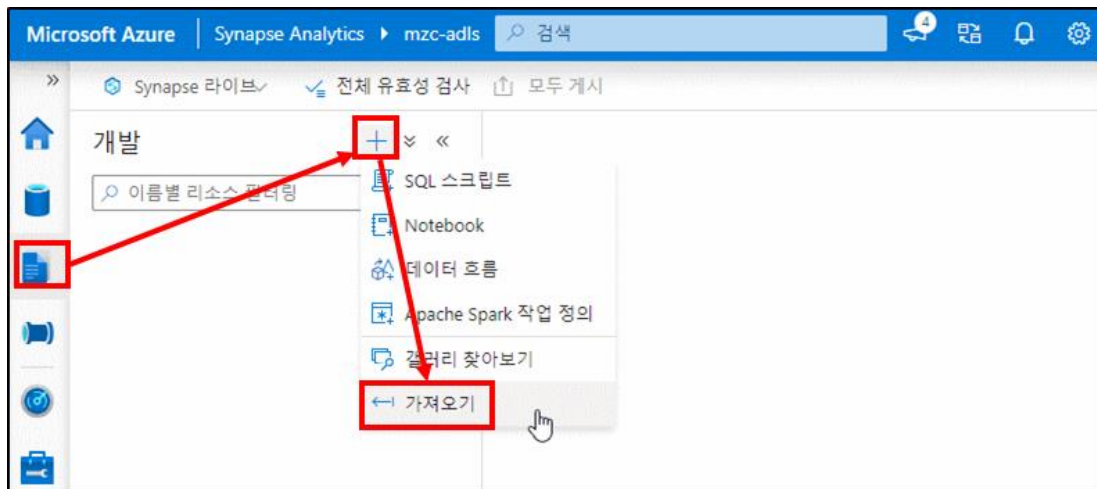
Task 1 : 작업 Script 불러오기

1. Azure Synapse 작업 영역의 개요 항목에서 작업 영역 웹 URL의 Link를 클릭하여 Synapse Studio를 실행합니다.

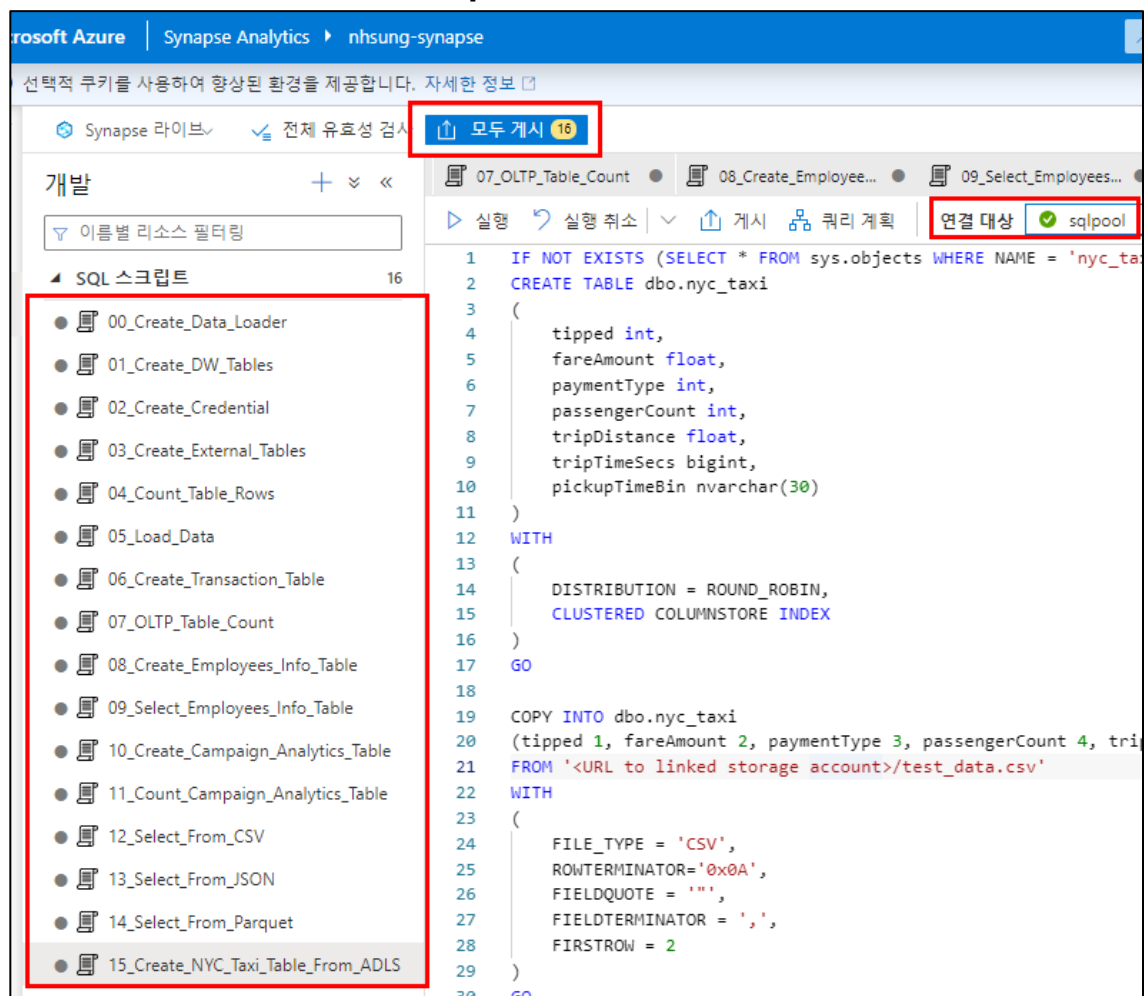
The screenshot displays the Azure portal interface for a Synapse workspace named 'mzc-adls'. The left-hand navigation pane shows the '개요' (Overview) tab selected. A red arrow originates from this tab and points to the '작업 영역 웹 URL' (Workspace Web URL) field in the '기본 정보' (Basic Information) section on the right. This field contains the URL <https://web.azuresynapse.net/ko/?workspace=%2fsubscriptions/5555a06d-ebb5-409a-8704-ce990875a093/workspaces/mzc-adls>, which is enclosed in a red rectangular box. Below the main content area, there is a 'Synapse Analytics 작업 영역' (Synapse Analytics Workspace) section for 'mzc-adls', featuring a '새로 만들기' (Create New) button and three informational cards: '수집' (Collect), '탐색 및 분석' (Explore and Analyze), and '시각화' (Visualize).

2. 개발 Hub 탭을 선택한 후 + 를 눌러 가져오기를 선택합니다.

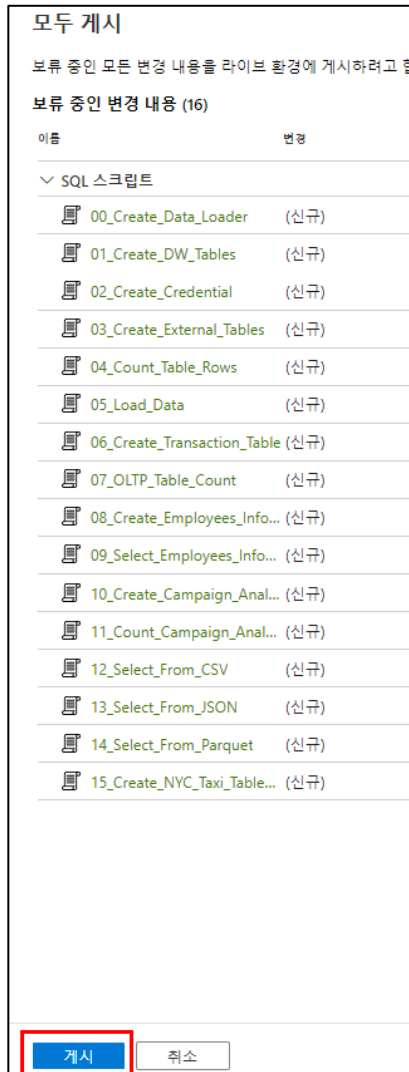
다운로드 받은 MZC Azure Synapse\LAB File\SQL Script 안의 **SQL Script** 들을 선택하여 불러옵니다. (16개 SQL 스크립트)



3. SQL 스크립트가 정상적으로 Import 되었는지 확인 합니다.



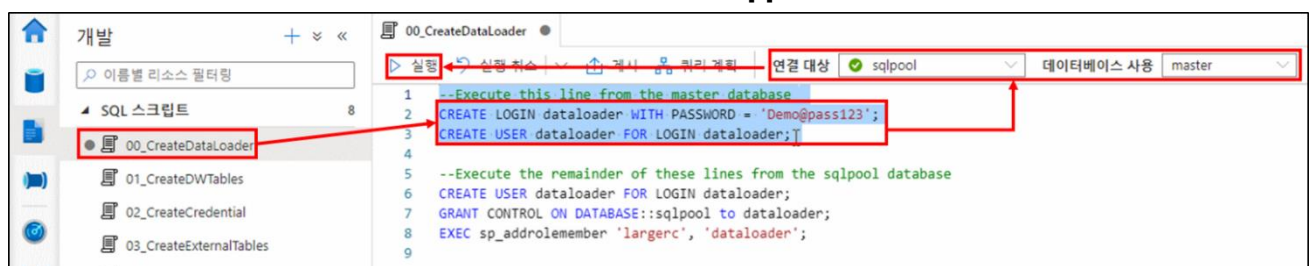
4. 모두 게시 → 게시를 선택하여 가져온 SQL 스크립트들을 자신의 Studio 에 저장하도록 합니다.

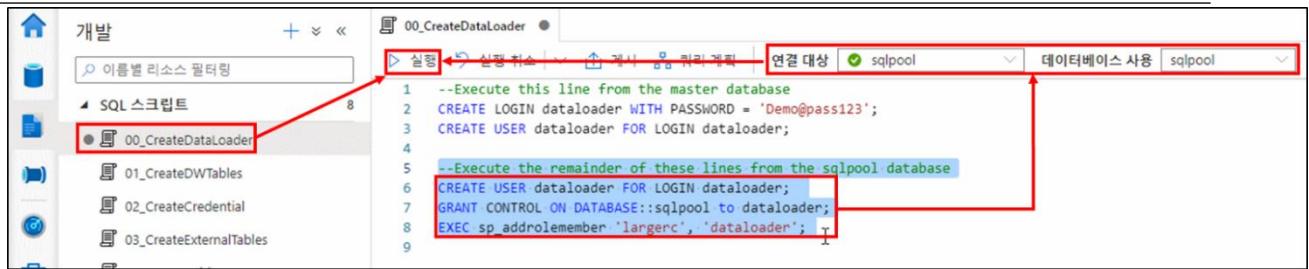


Task 2 : 테이블 생성, 데이터 Load, 데이터 확인

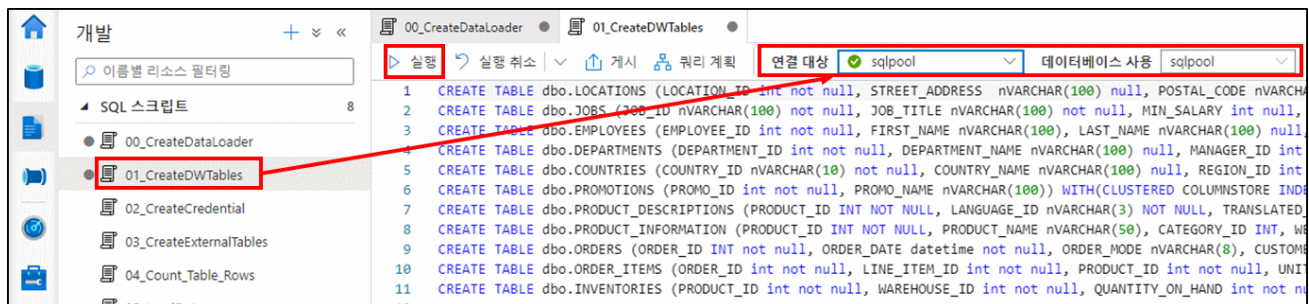
1. Data Loader 계정 생성을 위해 **00_CreateDataLoader** 스크립트를 실행합니다.

먼저 **Connect to** 에는 **sqlpool** 을 선택하고, 위의 2 줄은 데이터베이스를 **master** 선택하여 실행하고 아래 2 줄은 데이터베이스를 **sqlpool** 을 선택하여 실행합니다

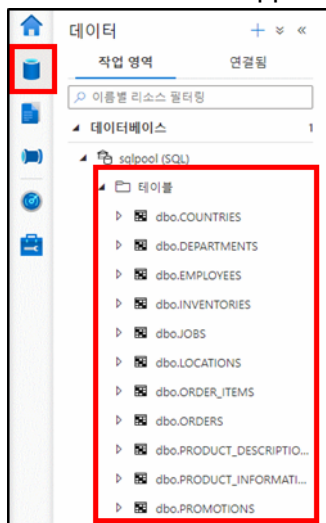




2. Oracle Server VM 의 DW 테이블들의 데이터 이관을 위해 Oracle DW 테이블 스키마에 맞춰 SQL Pool 안에 테이블들을 생성할 것입니다. **01_CreateDWTables** 스크립트를 수행합니다.



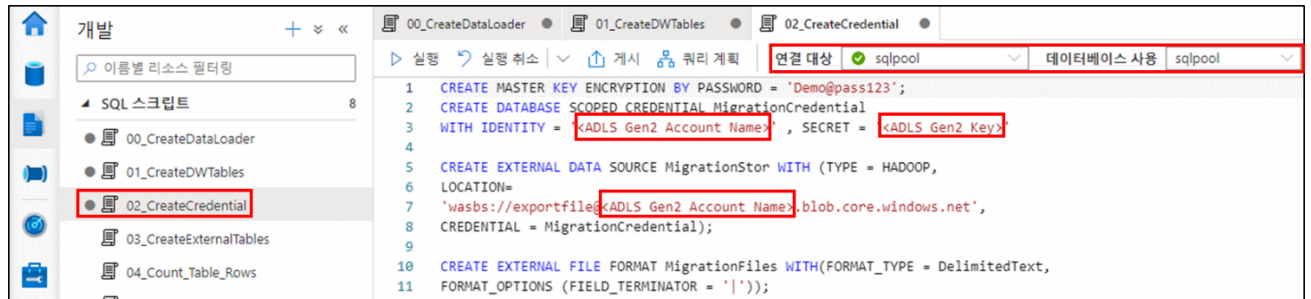
3. SQL Pool 안에 테이블들이 정상적으로 생성되었는지 확인하기 위해 **데이터 Hub** 탭을 선택하고 sqlpool 안의 **테이블 List** 를 확인합니다.



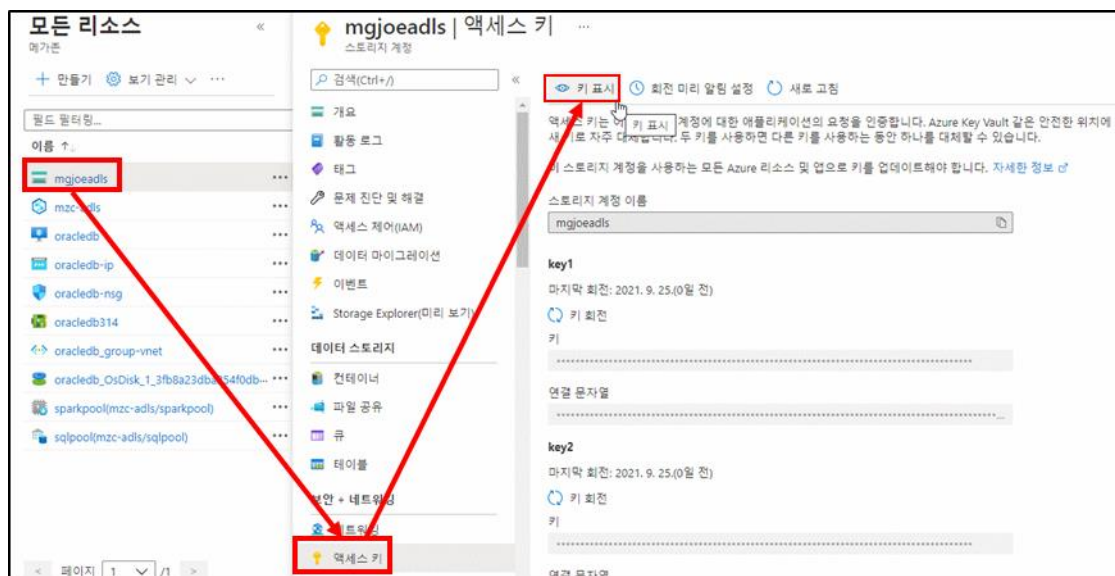
4. **02_CreateCredential** 스크립트를 수행 절차. (주의 : 해당 스크립트를 바로 실행하지 마세요.)

- Sqlpool 에서 exportfile 컨테이너에 접근하기 위한 Master Key, Credential, Data Source, File Format 을 정의하는 과정입니다.
- Oracle DW 테이블에서 Export 받은 파일들을 ADLS Gen2 의 'exportfile' 컨테이너 안에 올려 놓았습니다.

- 개발 Hub 로 돌아와서 **02_CreateCredential** 스크립트를 선택합니다.
- **02_CreateCredential** 스크립트 내용 중 <ADLS Gen2 Account Name>, <ADLS Gen2 Key> 를 자신의 환경에 맞게 내용을 바꾼 후 스크립트를 실행해야 합니다.

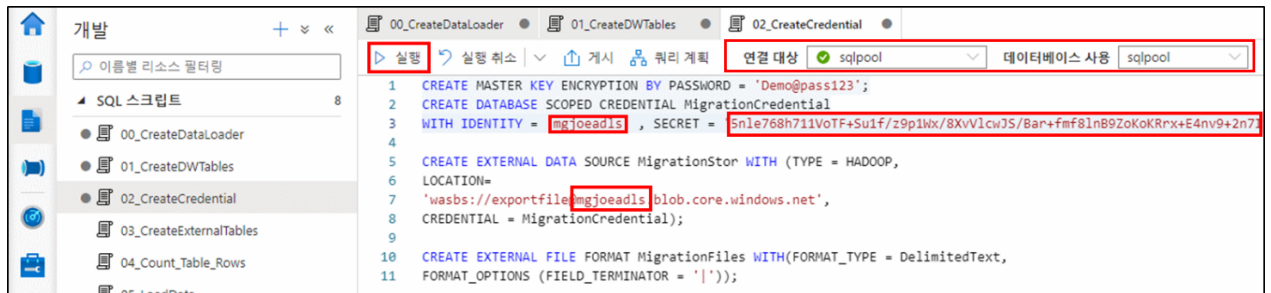


- <ADLS Gen2 Account Name>, <ADLS Gen2 Key> 를 확인하기 위하여 Azure Portal 의 **Storage Account** 를 선택 후 **액세스 키** 탭을 선택합니다.

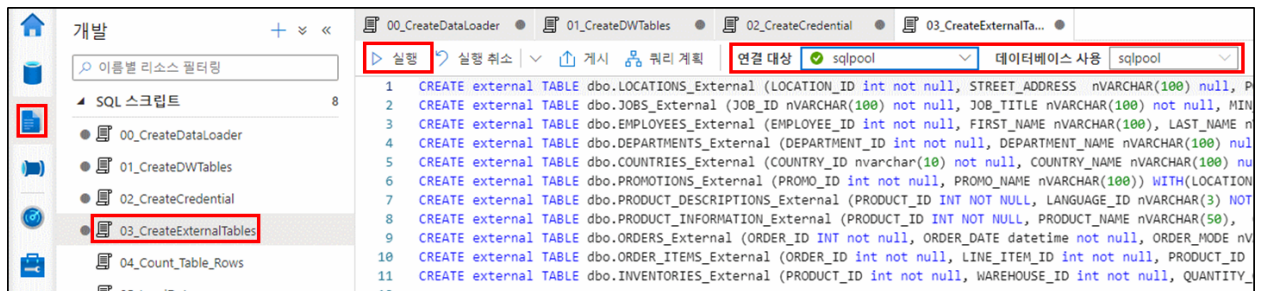


- 액세스 키 탭에서 키 표시를 눌러 Key 값들을 확인 및 복사하여

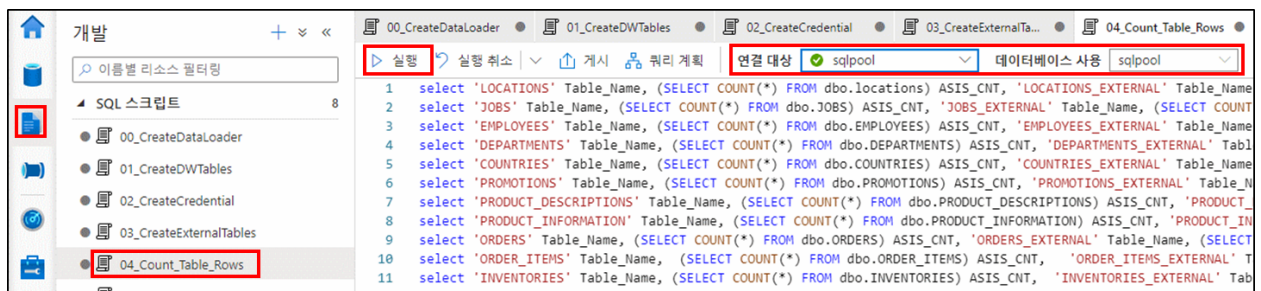
02_CreateCredential 스크립트의 <ADLS Gen2 Account Name>, <ADLS Gen2 Key> 값을 수정하여 스크립트를 실행합니다.



5. 다음으로, **03_CreateExternalTables** 스크립트를 실행하여 Export 받은 파일들을 직접 Link 하는 External 테이블들을 생성합니다.



6. **04_Count_Table_Rows** 스크립트를 실행하여 sqlpool의 테이블과 External 테이블의 건수를 비교합니다.



- **04_Count_Table_Rows** 스크립트 실행 결과입니다.
- SQL Pool에 생성된 테이블은 아직 Data가 Load되지 않아 0건으로 나오는 반면 External 테이블은 ADLS Gen2의 파일에서 바로 Data를 Load하기 때문에 0건 이상의 row가 나옵니다.

결과 메시지			
<div> <div>보기</div> <div>테이블</div> <div>차트</div> <div>결과 내보내기</div> </div>			
<div> <div>검색</div> </div>			
Table_Name	ASIS_CNT	Table_Name	TOBE_CNT
LOCATIONS	0	LOCATIONS_EXTERNAL	23
JOBS	0	JOBS_EXTERNAL	19
EMPLOYEES	0	EMPLOYEES_EXTERNAL	107
DEPARTMENTS	0	DEPARTMENTS_EXTERNAL	27
COUNTRIES	0	COUNTRIES_EXTERNAL	25
PROMOTIONS	0	PROMOTIONS_EXTERNAL	2
PRODUCT_DESCRIPTIONS	0	PRODUCT_DESCRIPTIONS_EXTER...	8640
PRODUCT_INFORMATION	0	PRODUCT_INFORMATION_EXTER...	288
ORDERS	0	ORDERS_EXTERNAL	105
ORDER_ITEMS	0	ORDER_ITEMS_EXTERNAL	665
INVENTORIES	0	INVENTORIES_EXTERNAL	1112

7. **05_Load_Data** 스크립트를 실행하여 External 테이블의 Data 를 테이블로 Insert 합니다. 스크립트 실행 후 하단의 결과에 SQL Pool 의 테이블과 External 테이블의 Count 가 일치하는지 비교해봅니다.

개발

이름별 리소스 필터링

SQL 스크립트 16

00_Create_Data_Loader

01_Create_DW_Tables

02_Create_Credential

03_Create_External_Tables

04_Count_Table_Rows

05_Load_Data

06_Create_Transaction_Table

07_OLTP_Table_Count

08_Create_Employees_Info_Table

09_Select_Employees_Info_Table

10_Create_Campaign_Analytics_Table

11_Count_Campaign_Analytics_Table

12_Select_From_CSV

13_Select_From_JSON

14_Select_From_Parquet

15_Create_NYC_Taxi_Table_From_ADLS

00_Create_Data_Lo...

01_Create_DW_Tables

02_Create_Credential

03_Create_External_T...

04_Count_Table_Rows

실행 실행 취소

계시

쿼리 계획

연결 대상 sqlpool

데이터베이스 사용 sqlpool

```

1 insert into dbo.locations select LOCATION_ID,STREET_ADDRESS,POSTAL_CODE,CITY,STATE_PROVINCE,COUNTRY_ID from
2 insert into dbo.jobs select JOB_ID,JOB_TITLE,MIN_SALARY,MAX_SALARY from dbo.JOBS_External;
3 insert into dbo.employees select EMPLOYEE_ID,FIRST_NAME,LAST_NAME,EMAIL,PHONE_NUMBER,HIRE_DATE,JOB_ID,SALARY
4 insert into dbo.departments select DEPARTMENT_ID,DEPARTMENT_NAME,MANAGER_ID,LOCATION_ID from dbo.DEPARTMENTS
5 insert into dbo.countries select COUNTRY_ID,COUNTRY_NAME,REGION_ID from dbo.COUNTRIES_External;
6 insert into dbo.PROMOTIONS select PROMO_ID,PROMO_NAME from dbo.PROMOTIONS_External;
7 insert into dbo.PRODUCT_DESCRIPTIONS select PRODUCT_ID,LANGUAGE_ID,TRANSLATED_NAME,TRANSLATED_DESCRIPTION fr
8 insert into dbo.PRODUCT_INFORMATION select PRODUCT_ID,PRODUCT_NAME,CATEGORY_ID,WEIGHT_CLASS,WARRANTY_PERIOD,
9 insert into dbo.ORDERS select ORDER_ID,ORDER_DATE,ORDER_MODE,CUSTOMER_ID,ORDER_STATUS,ORDER_TOTAL,SALES_REP
10 insert into dbo.ORDER_ITEMS select ORDER_ID,LINE_ITEM_ID,PRODUCT_ID,UNIT_PRICE,QUANTITY from dbo.ORDER_ITEMS
11 insert into dbo.INVENTORIES select PRODUCT_ID,WAREHOUSE_ID,QUANTITY_ON_HAND from dbo.INVENTORIES_External;
12
13 select count(*) from dbo.LOCATIONS UNION ALL
14 select count(*) from dbo.JOBS UNION ALL
15 select count(*) from dbo.EMPLOYEES UNION ALL
16 select count(*) from dbo.DEPARTMENTS UNION ALL
17 select count(*) from dbo.COUNTRIES UNION ALL
18 select count(*) from dbo.PROMOTIONS UNION ALL
19 select count(*) from dbo.PRODUCT_DESCRIPTIONS UNION ALL
20 select count(*) from dbo.PRODUCT_INFORMATION UNION ALL
21 select count(*) from dbo.ORDERS UNION ALL

```

결과 메시지

보기

테이블

차트

결과 내보내기

검색

(열 이름 없음)

ASIS_CNT

(열 이름 없음)

TOBE_CNT

LOCATIONS

23

LOCATIONS_EXTERNAL

23

JOBS

19

JOBS_EXTERNAL

19

EMPLOYEES

107

EMPLOYEES_EXTERNAL

107

DEPARTMENTS

27

DEPARTMENTS_EXTERNAL

27

COUNTRIES

25

COUNTRIES_EXTERNAL

25

PROMOTIONS

2

PROMOTIONS_EXTERNAL

2