



Azure Service

ASO Hands on Lab

Contents

1. 사전 준비	2
1.1. 시나리오 개요	2
1.2. 실습대상	2
1.3. Azure 서비스 및 관련 제품	2
1.4. 필수 조건	2
2. 학습 목표	3
2.1. 솔루션 아키텍처	3
2.2. 대상 환경 정보	3
3. Exercise 1: Configure Azure Services.....	4
3.1. Task 1: Create an Azure Storage Account.....	4
3.2. Task 2: Create an Azure SQL Data Warehouse	8
3.3. Task 3: Prepare Azure SQL Data Warehouse	11
3.4. Task 4: Create an Azure Data Factory v2	15
4. Exercise 2: Orchestrate with Azure Data Factory.....	17
4.1. Task 1: Amazon S3 to Azure Gen2.....	17
4.2. Task 2: MySQL to Azure Synapse Analytics.....	26
4.3. Task 3: Azure ML Execute	43
4.4. Task 4: PreProcessing and DataExport	45
4.5. Task 5: Send Email	47
5. Exercise 3: Visualize data with Power BI Desktop	50
5.1. Task 1: Install Power BI Desktop	50
5.2. Task 2: Query data with Power BI Desktop	53

1. 사전 준비

1.1. 시나리오 개요

1.2. 실습대상

- Database Administrators
- Data Architects
- Data Science

1.3. Azure 서비스 및 관련 제품

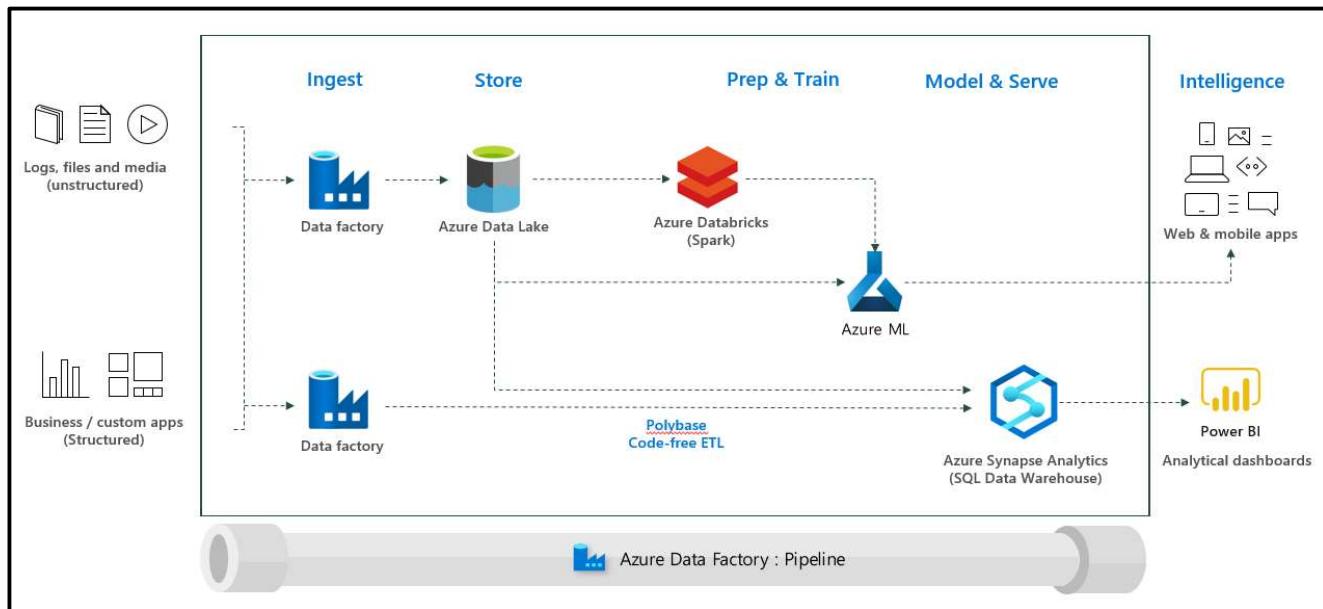
- Azure Data Factory v2
- Azure Storage
- Azure SQL Data Warehouse
- Azure Databricks Service
- Machine Learning
- Power BI

1.4. 필수 조건

- Azure Subscription

2. 학습 목표

2.1. 솔루션 아키텍처



2.2. 대상 환경 정보

이 연습에서는 이 실습을 위한 소스 환경을 배포합니다.

1. Azure Portal (<https://portal.azure.com>)에서 이 랩에 사용할 구독으로 로그인 했는지 확인합니다.

3. Exercise 1: Configure Azure Services

이 실습에서는 Azure Storage Account, Azure SQL Data Warehouse 및 Azure Data Factory V2를 생성하고 구성합니다. 이러한 서비스를 사용하여 기존 Data Warehouse 및 Data File을 Azure로 마이그레이션 할 수 있습니다.

3.1. Task 1: Create an Azure Storage Account

1. Azure 포탈로 이동하여 <https://portal.azure.com/>에서 로그인합니다.
2. 홈 화면 상단의 리소스 만들기를 선택하여, **storage**를 검색하고 Storage 계정을 선택합니다.

The screenshot shows the Azure Marketplace search results for 'storage'. A search bar at the top contains the text 'storage'. Below it, a list of services is shown. On the left, there's a sidebar with categories like '시작', 'AI + 기계 학습', '분석', '블록체인', '컴퓨팅', and '컨테이너'. The main area displays two service cards. The first card is for 'Data Lake Storage Gen1' by Microsoft, described as '빅 데이터 분석 작업을 위한 하이퍼 스케일(hyper-scale) 리포지토리'. The second card, which is highlighted with a red border, is for 'Storage 계정 - Blob, File, Table, Queue' by Microsoft, described as '안정적이고 경제적인 클라우드 저장소에 Blob, 테이블, 큐 및 파일을 사용할 수 있습니다.' Both cards have a blue heart icon at the bottom right.

3. 저장소 계정 블레이드에서 만들기를 클릭합니다. 다음 정보를 입력합니다.

- ✓ 구독: **본인의 구독 선택**
- ✓ 리소스 그룹: 이전에 생성한 리소스 선택 -> **ASO_EduXXXX**
- ✓ 스토리지 계정 이름: **asostorageXXXX**
- ✓ 위치: **한국 중부(Korea Central)**.
- ✓ 성능: **표준**
- ✓ 계정 종류: **StorageV2 (범용 v2)**
- ✓ Replication: **LRS(로컬 중복 스토리지)**
- ✓ 액세스 계층: **핫**

스토리지 계정 만들기

기본 사항 네트워킹 고급 태그 검토 + 만들기

Azure Storage는 가용성, 보안, 내구성, 확장성 및 중복성이 뛰어난 클라우드 스토리지를 제공하는 Microsoft 관리 서비스입니다. Azure Storage는 Azure Blob(개체), Azure Data Lake Storage Gen2, Azure Files, Azure 큐 및 Azure 테이블을 포함합니다. 스토리지 계정의 비용은 사용량 및 아래에서 선택한 옵션에 따라 다릅니다. [Azure Storage 계정에 대한 자세한 정보](#)

프로젝트 정보

배포된 리소스와 비용을 관리할 구독을 선택합니다. 풀더 같은 리소스 그룹을 사용하여 모든 리소스를 정리 및 관리합니다.

구독 *

Azure subscription 1

리소스 그룹 *

ASO_Edu0000

[새로 만들기](#)

인스턴스 정보

기본 배포 모델은 최신 Azure 기능을 지원하는 Resource Manager입니다. 대신 클래식 배포 모델을 사용하여 배포하도록 선택할 수 있습니다. [클래식 배포 모델 선택](#)

스토리지 계정 이름 * ⓘ

asostorage0000

위치 *

(아시아 태평양)한국 중부

성능 ⓘ

표준 프리미엄

계정 종류 ⓘ

StorageV2(범용 v2)

복제 ⓘ

LRS(로컬 중복 스토리지)

액세스 계층(기본값) ⓘ

블 핫

[검토 + 만들기](#)

< 이전

다음: 네트워킹 >

- 상단의 고급을 클릭하고 Data Lake Storage Gen2를 사용으로 변경합니다.

스토리지 계정 만들기

기본 사항 네트워킹 고급 태그 검토 + 만들기

보안

보안 전송 필요 ⓘ 사용 안 함 사용

Azure Files

대용량 파일 공유 ⓘ 사용 안 함 사용

데이터 보호

Blob 일시 삭제 ⓘ 사용 안 함 사용

데이터 보호 및 계층 구조 네임스페이스를 동시에 사용하도록 설정할 수 없습니다.

Data Lake Storage Gen2

계층 구조 네임스페이스 ⓘ 사용 안 함 사용

NFS v3 ⓘ

사용 안 함 사용

등록은 현재 구독별로 NFS v3 기능을 활용하는 데 필요합니다. [NFS v3 등록](#)

[검토 + 만들기](#)

< 이전

다음: 태그 >

5. **검토+만들기**를 클릭하고 구성 선택 사항을 확인한 후 **만들기**를 선택합니다.

6. 새 스토리지 계정이 완성되면 리소스로 이동하여 **컨테이너**를 클릭합니다.

홈 > Microsoft.StorageAccount-20200220105915 - 개요 > asostorage0000

asostorage0000
스트리지 계정

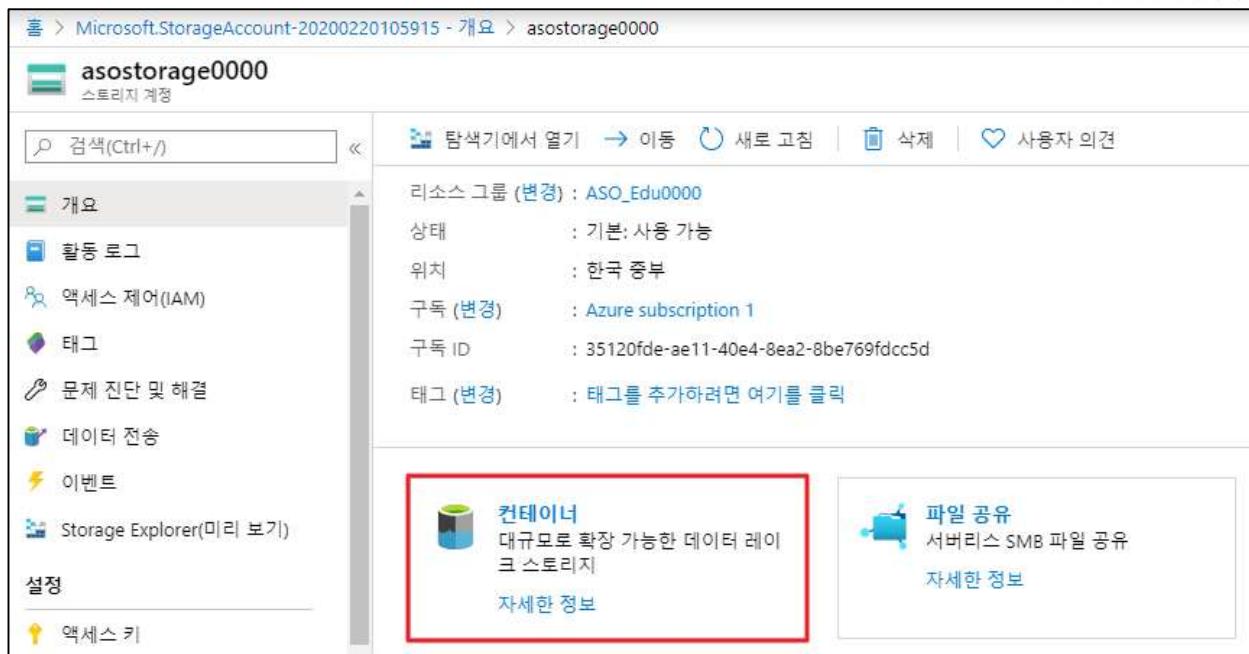
검색(Ctrl+ /) <>

리소스 그룹 (변경) : ASO_Edu0000
상태 : 기본: 사용 가능
위치 : 한국 중부
구독 (변경) : Azure subscription 1
구독 ID : 35120fde-ae11-40e4-8ea2-8be769fdcc5d
태그 (변경) : 태그를 추가하려면 여기를 클릭

개요 활동 로그 액세스 제어(IAM) 태그 문제 진단 및 해결 데이터 전송 이벤트 Storage Explorer(미리 보기) 설정 액세스 키

컨테이너
대규모로 확장 가능한 데이터 레이크 스토리지
자세한 정보

파일 공유
서비스 SMB 파일 공유
자세한 정보



7. +컨테이너를 클릭한 후 demo를 입력하고 확인 버튼을 클릭합니다.

홈 > 리소스 그룹 > ASO_Edu0000 > asostorage0000 - 컨테이너

asostorage0000 - 컨테이너
스트리지 계정

검색(Ctrl+ /) <>

+ 컨테이너 액세스 수준 변경 새로 고침 삭제

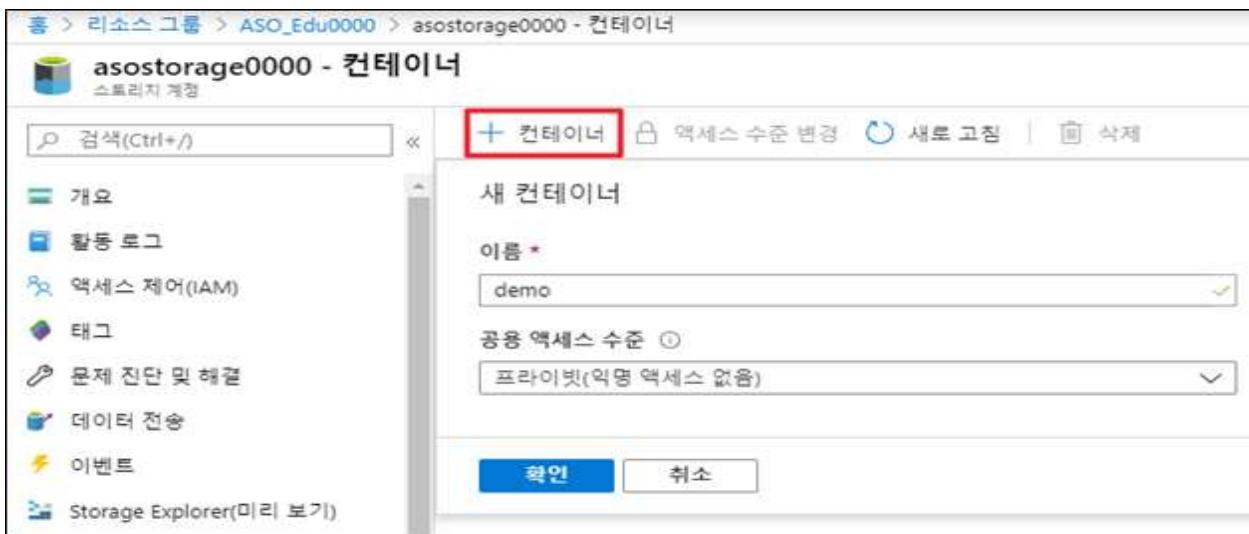
개요 활동 로그 액세스 제어(IAM) 태그 문제 진단 및 해결 데이터 전송 이벤트 Storage Explorer(미리 보기)

새 컨테이너

이름 * : demo

공용 액세스 수준 : 프라이빗(익명 액세스 없음)

확인 취소



8. 위와 같은 방식으로 +컨테이너를 클릭 하여 structured를 입력하고 확인 버튼을 클릭합니다.

홈 > asostorage0000 - 컨테이너

asostorage0000 - 컨테이너
스트리지 계정

검색(Ctrl+ /) <>

+ 컨테이너 액세스 수준 변경 새로 고침 삭제

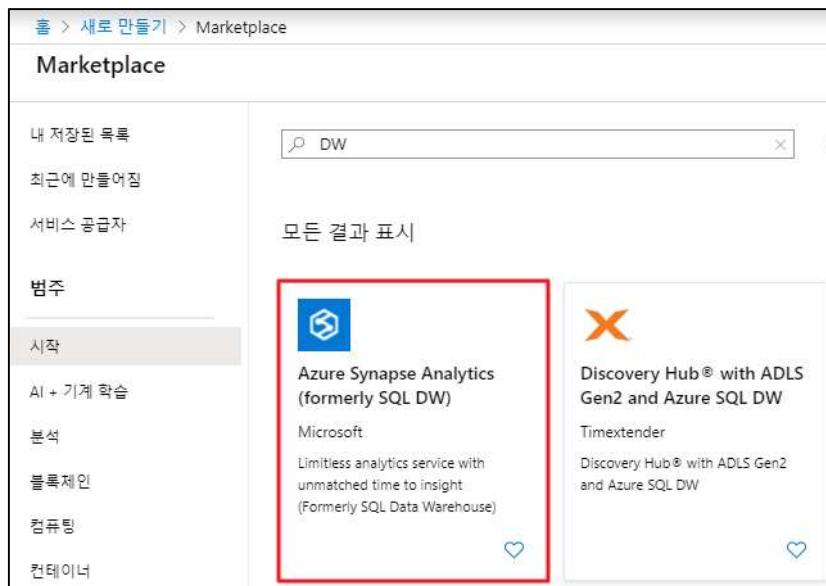
접두사로 컨테이너 검색

이름	마지막으로 수정한 날...	공용 액세스 ...	임대 단계
demo	2020. 2. 20. 오전 11:18...	프라이빗	사용 가능
structured	2020. 2. 20. 오후 1:41:18	프라이빗	사용 가능



3.2. Task 2: Create an Azure SQL Data Warehouse

1. Azure 포탈로 이동하여 <https://portal.azure.com/>에서 로그인합니다.
2. 왼쪽 메뉴의 리소스 만들기를 선택하여, DW 입력하여 검색한 후 Azure Synapse Analytics를 선택합니다.



3. Azure Synapse Analytics 블레이드에서 만들기를 클릭하고 아래 내용을 입력합니다.

- ✓ 구독: 본인의 구독 선택.
- ✓ 리소스 그룹: 이전에 생성한 리소스 선택 -> ASO_EduXXXX
- ✓ 데이터 웨어하우스 이름: CohoDWXXXX

홈 > 새로 만들기 > Azure Synapse Analytics (formerly SQL DW) > Azure Synapse Analytics

Azure Synapse Analytics

Microsoft

Azure Synapse Analytics(이전 이름 Azure SQL Data Warehouse)를 시작합니다. [자세한 정보](#)

[기본](#) * [추가 설정 *](#) [태그](#) [검토 + 만들기](#)

원하는 구성으로 SQL 데이터 웨어하우스를 만듭니다. [기본] 탭을 완료한 후 [검토] + [만들기]로 가서 적절한 기본값으로
프로비전하거나, 각 탭을 방문하여 사용자 지정합니다. [자세히 보기](#)

프로젝트 세부 정보

배포된 리소스와 비용을 관리할 구독을 선택합니다. 폴더 같은 리소스 그룹을 사용하여 모든 리소스를 정리 및 관리합니다.

구독 * ⓘ

리소스 그룹 * ⓘ [새로 만들기](#)

데이터 웨어하우스 세부 정보

논리 서버를 선택하고 성능 수준을 구성하는 것을 포함하여 이 데이터 웨어하우스의 필수 설정을 입력합니다.

데이터 웨어하우스 이름 *

서버 * ⓘ [새로 만들기](#)

성능 수준 * ⓘ [성능 수준 선택](#)

[검토 + 만들기](#) [다음: 추가 설정 >](#)

4. 서버에서 **새로 만들기**를 클릭하여 아래 내용을 입력 후 **확인** 버튼을 클릭합니다.

- ✓ 서버 이름: **asodwserverXXXX** (DW 계정 명)
- ✓ 서버 관리자 로그인: **demouser**
- ✓ 암호: **Demo@pass123**
- ✓ 위치: **한국 중부(Korea Central)**
- ✓ Azure 서비스의 서버 액세스 허용: **체크함**

새 서버

Microsoft

서버 이름 *

 .database.windows.net

서버 관리자 로그인 *

암호 *

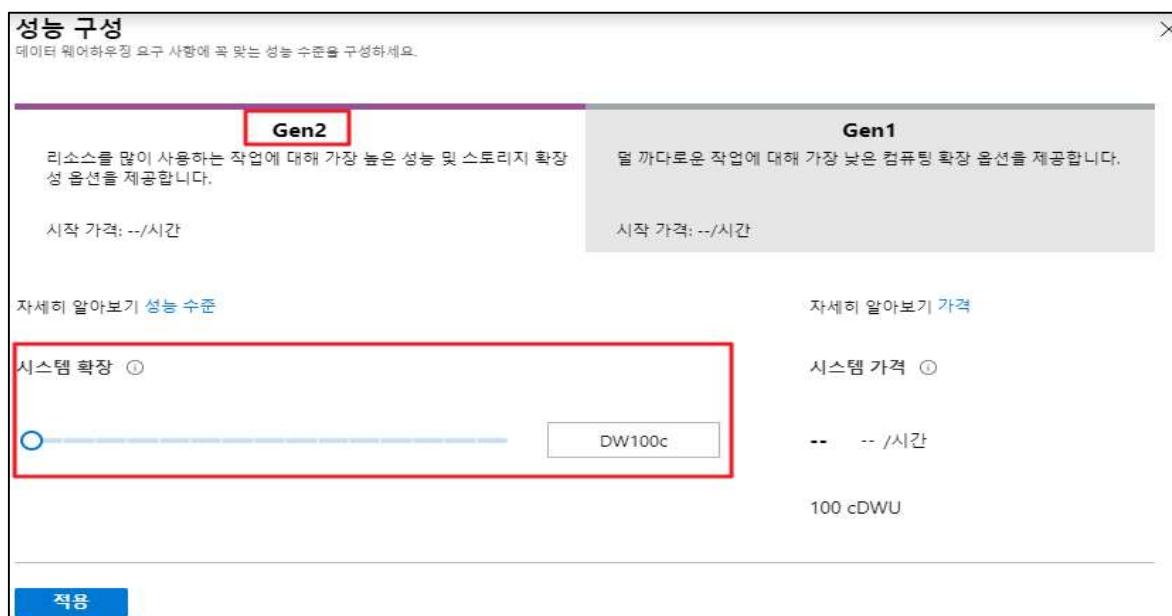
암호 확인 *

위치 *

Azure 서비스의 서버 액세스 허용 ⓘ

확인

- 성능 수준 선택을 클릭하고 Gen2 탭을 선택한 다음 성능을 100cDWU로 설정하고 적용을 클릭합니다.



- SQL Data Warehouse 블레이드에서 검토+만들기를 클릭한 후 만들기를 클릭합니다.

3.3. Task 3: Prepare Azure SQL Data Warehouse

1. Azure 포탈로 이동하여 <https://portal.azure.com>에서 로그인합니다.
2. 리소스 그룹으로 이동하여 **Data Warehouse CohoDW**를 선택합니다..

The screenshot shows the Azure Resource Group 'ASO_Edu0000' interface. On the left, there's a sidebar with icons for Overview, Activity Log, IAM, Tags, Events, Quick Start, Resource Usage, Metrics, Policies, and Properties. The main area displays the 'Azure subscription 1' and its ID. It also shows a list of resources under the 'Data Factory(V2)', 'SQL Server', 'Storage Account', and 'SQL Data Warehouse' categories. The 'CohoDW(asodwserver0000/CohoDW)' resource is highlighted with a red border.

3. 왼쪽 보안 메뉴의 방화벽 및 가상 네트워크로 이동하여 **+클라이언트 IP 추가** 버튼 클릭 후 저장 버튼을 클릭하여 저장합니다.

The screenshot shows the 'CohoDW(asodwserver0000/CohoDW) - 방화벽 및 가상 네트워크' settings page. The left sidebar includes icons for Connection String, Properties, Encryption, Template Export, Security, Advanced Data Protection, and Firewall & Virtual Network. The main area has a '저장' (Save) button highlighted with a red box. Below it, there's a note about granting access to all databases from the specified IP. The 'Azure 서비스 및 리소스가 이 서버에 액세스할 수 있도록 허용' section shows a '설정' (Configure) button. A table lists a client IP address (223.62.175.102) with columns for Rule Name, Start IP, and End IP. The table also shows a row for 'ClientIPAddress_2020-2-20...' with IP addresses 223.62.203.89 and 223.62.203.89.

4. 왼쪽 일반 작업 메뉴에서 쿼리 편집기를 클릭 후 아래의 정보를 입력하여 로그인합니다.

- ✓ 로그인: demouser
- ✓ 암호: Demo@pass123

홈 > 리소스 그룹 > ASO_Edu0000 > CohoDW(asodwserver0000/CohoDW) - 쿼리 편집기(미리 보기)

CohoDW(asodwserver0000/CohoDW) - 쿼리 편집기(미리 보기)
SQL Data Warehouse

검색(Ctrl+ /) <> 로그인 새 쿼리 쿼리 열기 피드백

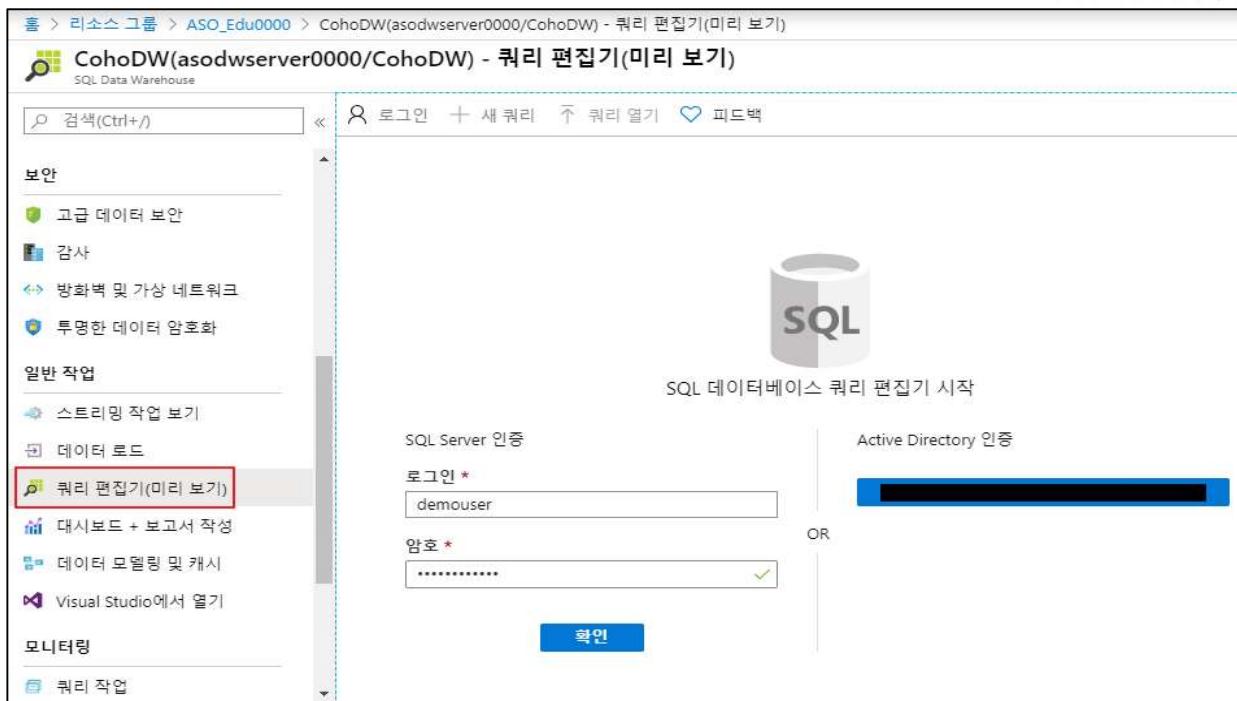
보안
고급 데이터 보안
감사
방화벽 및 가상 네트워크
투명한 데이터 암호화

일반 작업
스트리밍 작업 보기
데이터 로드
쿼리 편집기(미리 보기) (선택됨)
대시보드 + 보고서 작성
데이터 모델링 및 캐시
Visual Studio에서 열기

모니터링
쿼리 작업

SQL 데이터베이스 쿼리 편집기 시작

SQL Server 인증
로그인 * demouser
암호 *
OR
Active Directory 인증
[]
확인



5. 다음 쿼리를 실행하여 테이블을 생성합니다.

```

CREATE TABLE [dbo].[DimCustomer](
    [CustomerKey] int not null, [GeographyKey] int null,
    [CustomerAlternateKey] nvarchar(30) not null, [Title] nvarchar(16) null,
    [FirstName] nvarchar(100) null, [MiddleName] nvarchar(100) null,
    [LastName] nvarchar(100) null,
    [NameStyle] bit null,
    [BirthDate] date null,
    [MaritalStatus] nchar(2) null,
    [Suffix] nvarchar(20) null,
    [Gender] nvarchar(2) null,
    [EmailAddress] nvarchar(100) null,
    [YearlyIncome] money null,
    [TotalChildren] tinyint null,
    [NumberChildrenAtHome] tinyint null,
    [EnglishEducation] nvarchar(80) null,
    [SpanishEducation] nvarchar(80) null,
    [FrenchEducation] nvarchar(80) null,
    [EnglishOccupation] nvarchar(200) null,
    [SpanishOccupation] nvarchar(200) null,
    [FrenchOccupation] nvarchar(200) null,
    [HouseOwnerFlag] nchar(2) null,
    [NumberCarsOwned] tinyint null,
    [AddressLine1] nvarchar(240) null,
    [AddressLine2] nvarchar(240) null,
    [Phone] nvarchar(40) null,
    [DateFirstPurchase] date null,
    [CommuteDistance] nvarchar(30) null
) WITH(CLUSTERED COLUMNSTORE INDEX, DISTRIBUTION = ROUND_ROBIN);

CREATE TABLE [dbo].[DimGeography](
    [GeographyKey] int not null, [City] nvarchar(60) null,
    [StateProvinceCode] nvarchar(6) null,
    [StateProvinceName] nvarchar(100) null,
    [CountryRegionCode] nvarchar(6) null,
    [EnglishCountryRegionName] nvarchar(100) null,
    [SpanishCountryRegionName] nvarchar(100) null,
    [FrenchCountryRegionName] nvarchar(100) null,
    [PostalCode] nvarchar(30) null,
    [SalesTerritoryKey] int null
) WITH(CLUSTERED COLUMNSTORE INDEX, DISTRIBUTION = ROUND_ROBIN);

CREATE TABLE [dbo].[FactInternetSales](
    [ProductKey] int not null,
    [OrderDateKey] int not null,
    [DueDateKey] int not null,
    [ShipDateKey] int not null,
    [CustomerKey] int not null,
    [PromotionKey] int not null,
    [CurrencyKey] int not null,
    [SalesTerritoryKey] int not null,
    [SalesOrderNumber] nvarchar(40) not null,
    [SalesOrderLineNumber] tinyint not null,
    [RevisionNumber] tinyint not null,
    [OrderQuantity] smallint not null,
    [UnitPrice] money not null,
    [ExtendedAmount] money not null,
    [UnitPriceDiscountPct] float(8) not null,
    [DiscountAmount] float(8) not null,
    [ProductStandardCost] money not null,
    [TotalProductCost] money not null,
    [SalesAmount] money not null,
    [TaxAmt] money not null,
    [Freight] money not null,
    [CarrierTrackingNumber] nvarchar(50) null,
    [CustomerPONumber] nvarchar(50) null
) WITH(CLUSTERED COLUMNSTORE INDEX, DISTRIBUTION = ROUND_ROBIN)

```

```

1  CREATE TABLE [dbo].[DimCustomer](
2      [CustomerKey] int not null,
3      [GeographyKey] int null,
4      [CustomerAlternateKey] nvarchar(30) not null,
5      [Title] nvarchar(16) null,[FirstName] nvarchar(100) null,
6      [MiddleName] nvarchar(100) null,[LastName] nvarchar(100) null,
7      [NameStyle] bit null,[BirthDate] date null,
8      [MaritalStatus] nchar(2) null,
9      [Suffix] nvarchar(20) null,[Gender] nvarchar(2) null,
10     [EmailAddress] nvarchar(100) null,
11     [YearlyIncome] money null,
12     [TotalChildren] tinyint null,
13     [NumberChildrenAtHome] tinyint null,
14     [EmploymentStatus] nvarchar(100) null,
15     [EducationLevel] nvarchar(100) null,
16     [Occupation] nvarchar(100) null,
17     [LanguageId] int null,
18     [CultureId] int null,
19     [SalesTerritoryKey] int null
20 )

```

결과 메시지

쿼리 성공: Affected rows: 0

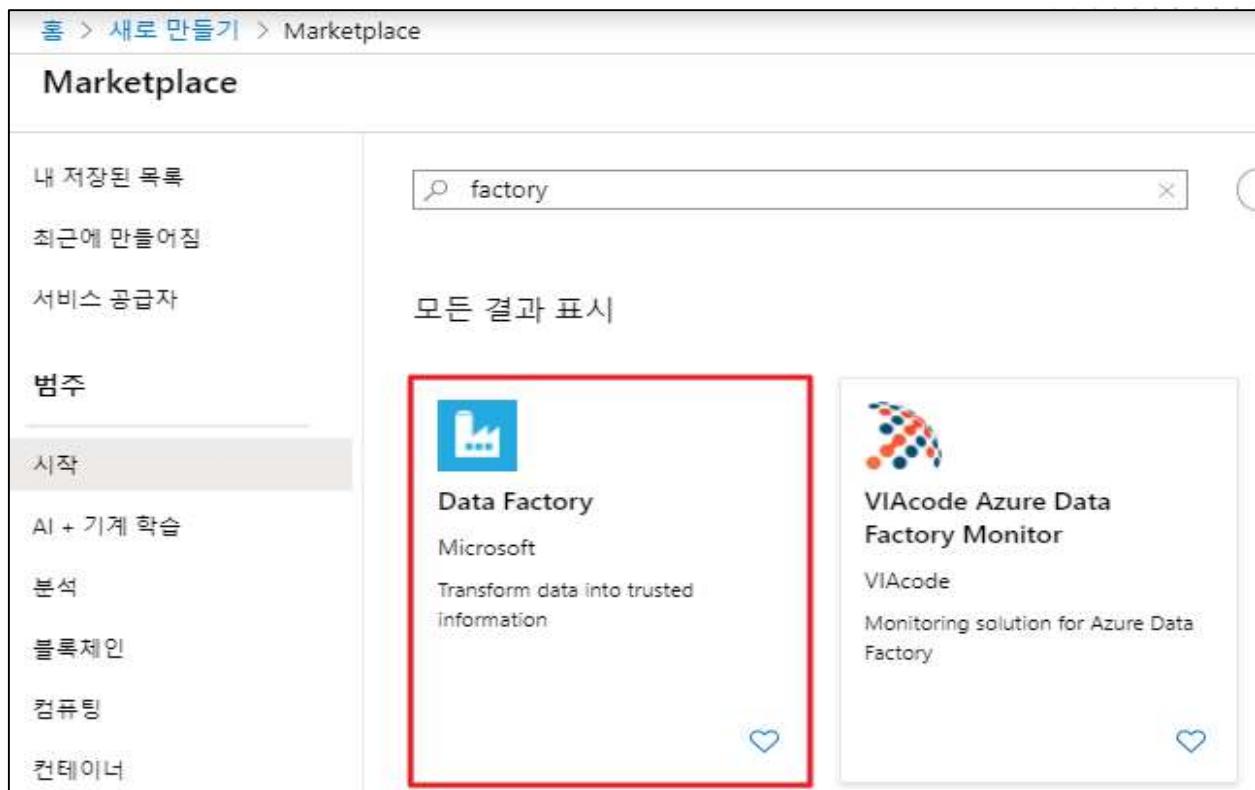
6. 다음 쿼리를 실행하여 테이블이 생성되었는지 확인합니다. 반환되는 행이 3개여야 합니다.

```
SELECT * FROM sys.tables
```

name	object_id	principal_id	schema_id	parent_object_id
DimCustomer	1986106116		1	0
DimGeography	1650104919		1	0
FactInternetSales	1666104976		1	0

3.4. Task 4: Create an Azure Data Factory v2

1. Azure 포탈로 이동하여 <https://portal.azure.com>에서 로그인합니다.
2. 리소스 만들기를 클릭하여 **Factory**를 입력, 검색하고 **Data Factory**를 선택합니다.



3. Data Factory 블레이드에서 만들기를 클릭한 후 다음 정보를 입력합니다.

- ✓ 이름: ASODatafactoryXXXX (XXXX는 사용자 별로 고유하게 입력)
- ✓ 버전: V2
- ✓ 구독: 본인의 구독 선택.
- ✓ 리소스 그룹: ASO_EduXXXX
- ✓ 위치: 한국 중부(Korea Central)
- ✓ Git 사용 미 체크

홈 > 리소스 그룹 > ASO_Edu0000 > 새로 만들기 > Data Factory > 새 data factory

새 data factory

이름 *

ASODatafactory0000

버전 ⓘ

V2

구독 *

Azure subscription 1

리소스 그룹 *

ASO_Edu0000

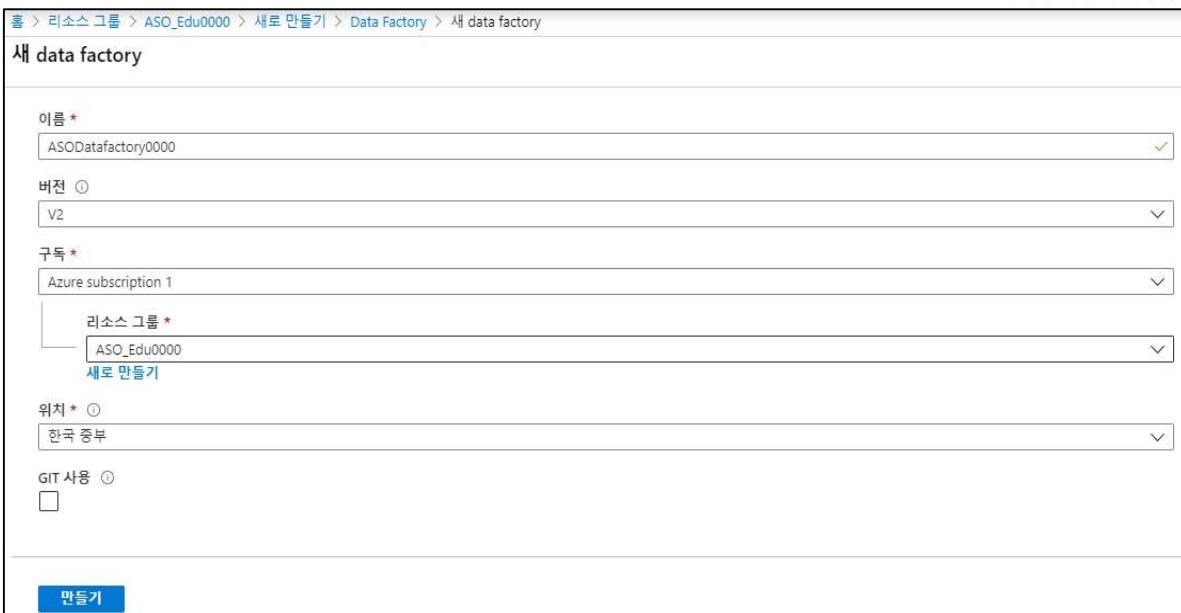
새로 만들기

위치 * ⓘ

한국 중부

GIT 사용 ⓘ

만들기



4. 만들기 버튼을 클릭합니다.

Data Factory 위치는 Data Factory의 메타데이터가 저장되고 파이프라인의 트리거가 시작되는 위치입니다. 한편, Data Factory는 다른 Azure 영역의 데이터 저장소 및 컴퓨팅 서비스에 액세스하여 데이터 저장소 간에 데이터를 이동하거나 컴퓨팅 서비스를 사용하여 데이터를 처리할 수 있습니다. 이 동작은 전 세계적으로 사용 가능한 통합 런타임을 통해 실현되어 데이터 컴플라이언스, 효율성 및 네트워크 송신 비용 절감을 보장합니다.

4. Exercise 2: Orchestrate with Azure Data Factory

이 실습에서는 Azure Data Factory의 Pipeline을 이용하여 특정 위치의 파일의 데이터를 이관하며, Azure Databricks, Azure Machine Learning 등 Azure의 다른 서비스들과 연결하여 전체적인 데이터의 흐름을 관리합니다.

4.1. Task 1: Amazon S3 to Azure Gen2

1. 리소스 그룹으로 이동하여 **Data Factory**를 선택 후 **작성 및 모니터링**을 클릭합니다.

The screenshot shows the Azure portal interface for managing a Data Factory resource named 'ASODatafactory0000'. The left sidebar lists various navigation options like '개요', '활동 로그', '액세스 제어(IAM)', '태그', '문제 진단 및 해결', '설정', '잠금', '일반', '속성', and '시작'. The main content area displays resource details such as '리소스 그룹 (변경) ASO_Edu0000', '형식 Data Factory(V2)', '상태 Succeeded', '시작하기 빠른 시작', '위치 한국 중부', '구독 (변경) Azure subscription 1', and '구독 ID 35120fde-ae11-40e4-8ea2-8be769fdcc5d'. At the bottom right, there are two buttons: '문서' and '작성 및 모니터링', with the latter being highlighted by a red box.

2. Data Factory 포털 홈 화면에서 **Create pipeline**을 선택합니다.

The screenshot shows the 'Let's get started' screen for Azure Data Factory. It features a central 'Create pipeline' button highlighted with a red box. Below it are four other options: 'Create data flow', 'Create pipeline from template', and 'Copy data', each accompanied by a small icon. The top navigation bar shows 'Microsoft Azure | Data Factory > ASODatafactory0000'.

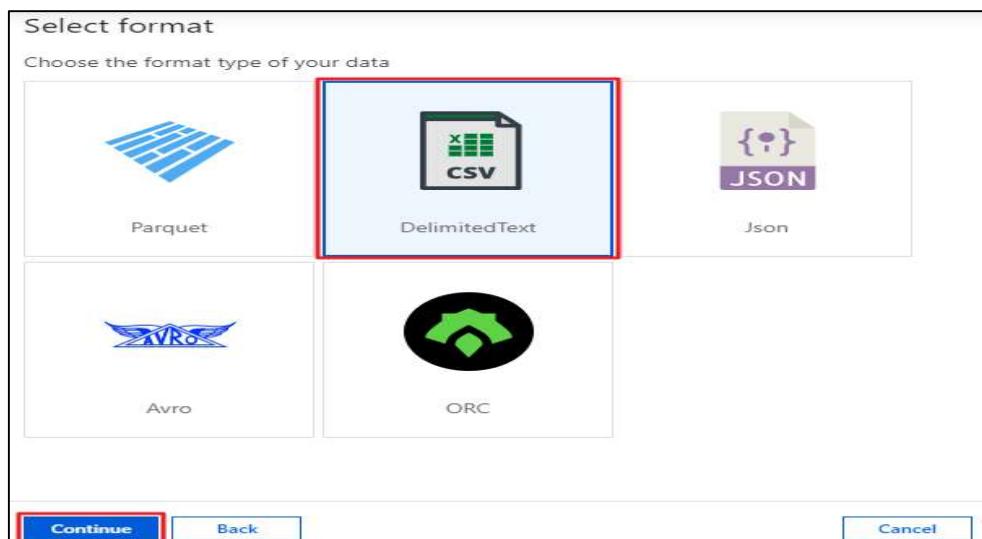
3. Activities > General > Lookup을 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.
 - ✓ Name: Check Amazon S3
4. Setting 탭으로 이동하여 Source dataset의 +New를 클릭합니다.



5. New dataset의 All 목록에서 Amazon S3 를 선택합니다.

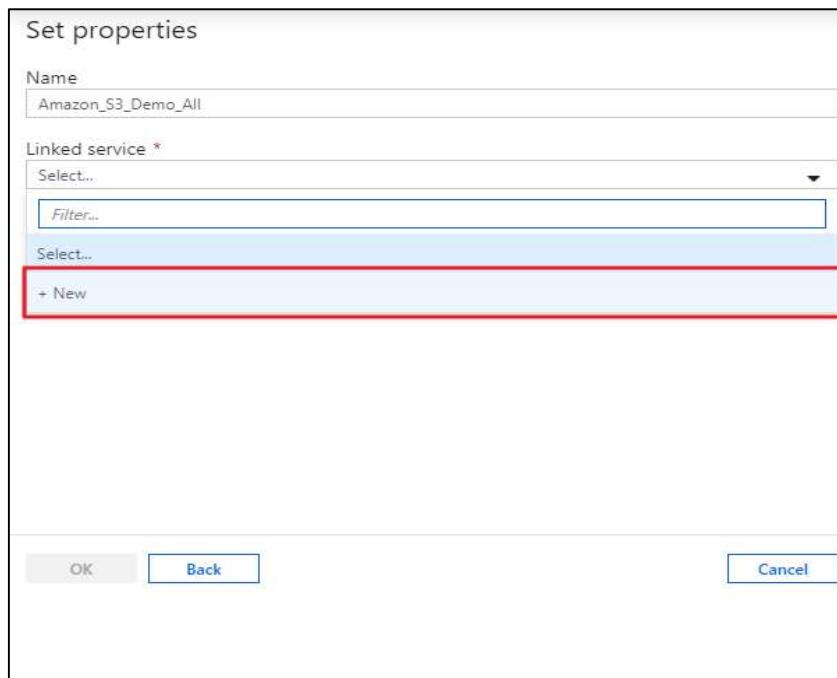


6. format은 DelimitedText를 선택합니다.



7. Name 항목 입력 후 Linked service 에서 +New를 선택합니다.

✓ Name : Amazon_S3_Demo_All



8. **New linked service (Amazon S3)** 블레이드가 열리면, 아래와 같이 정보를 입력하고 **Test Connection** 버튼을 클릭하여 Connection successful 메시지가 출력되면 **Create** 버튼을 클릭합니다.

- ✓ Name : AmazonS3
- ✓ Access Key ID : AKIAI54HTGVLRJ4TC7QQ
- ✓ Secret Access Key : OOr4sejC3ICrnT7Nuv4RaOezH7UDatbh4sl+ztgx

New linked service (Amazon S3)

If the identity you use to access the data store only has permission to subdirectory instead of the entire account, specify the path to test connection. Please make sure your self-hosted integration runtime is higher than version 4.0 if connecting via self-hosted integration runtime.

Name *
AmazonS3

Description

Connect via integration runtime *
AutoResolveIntegrationRuntime

Access Key ID *
AKIAI54HTGVLRJ4TC7QQ

Secret Access Key
 Azure Key Vault

Service URL

Test connection
 To linked service To file path

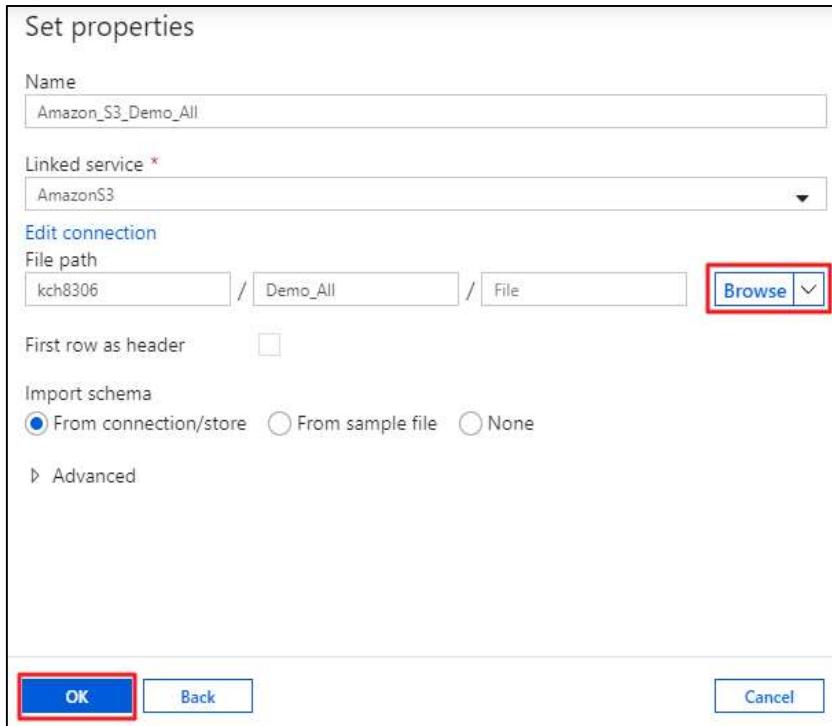
Annotations

Connection successful Test connection Create

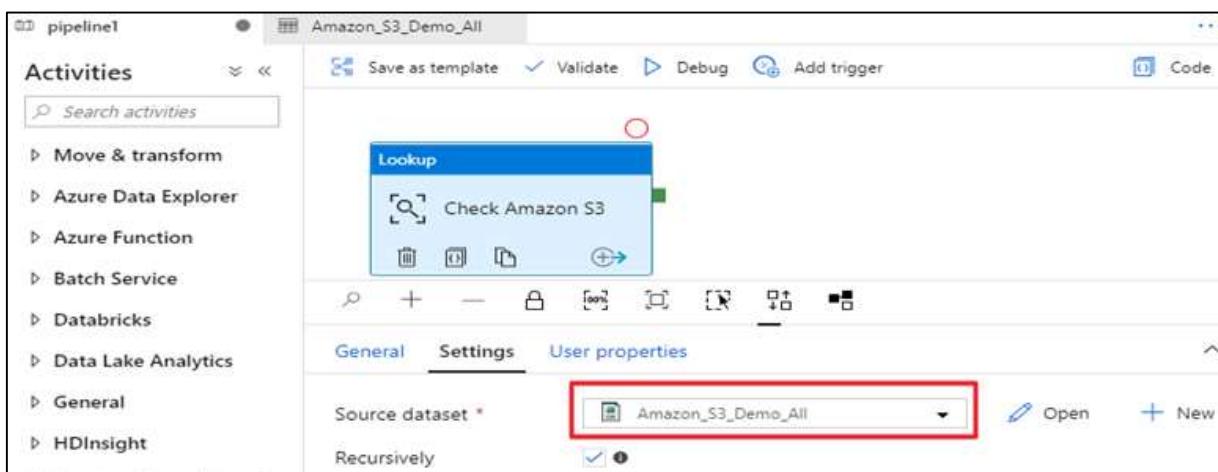
9. Set Properties 블레이드가 열리면 File path 항목의 Browse 버튼을 클릭하여 파일 경로를 kch8306

> Demo_All로 지정 후 OK 버튼을 클릭하여 완료합니다.

- ✓ Name : AmazonS3
- ✓ File path : kch8306 > Demo_All



10. Lookup 액티비티의 Setting 탭의 Source Dataset이 세팅된 것을 확인합니다.

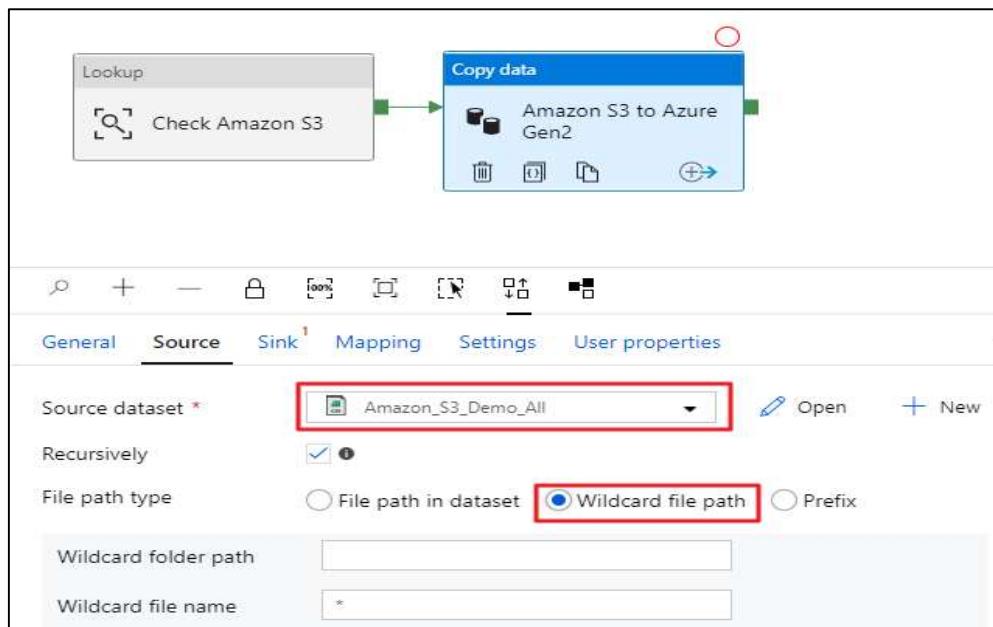


11. Activities > Move & transform > Copy data를 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.

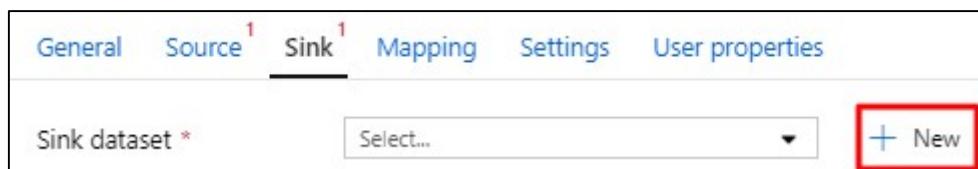
- ✓ Name: Amazon S3 to Azure Gen2

12. Lookup 액티비티와 Copy data 액티비티를 연결시킨 후 Source 탭으로 이동하여 다음을 입력합니다.

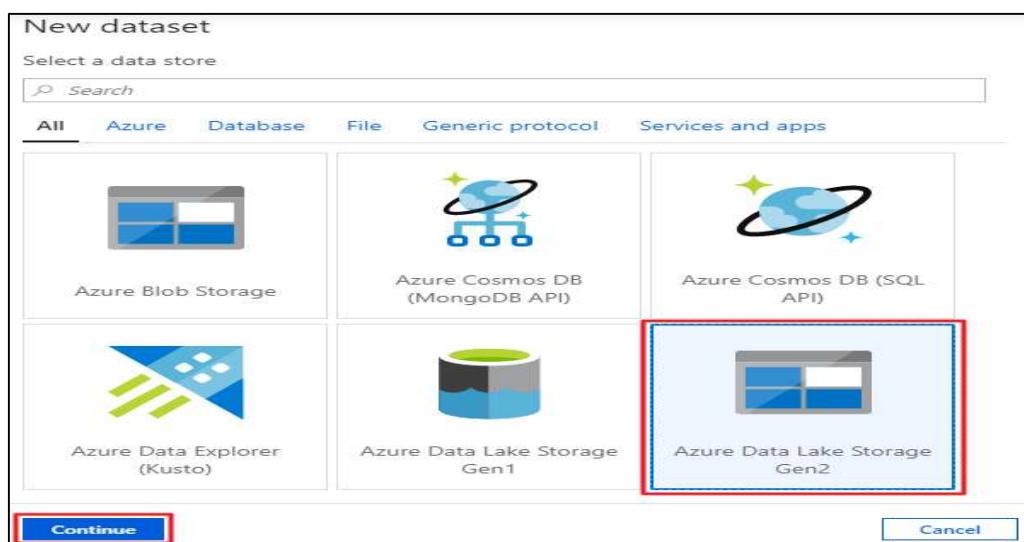
- ✓ Source dataset : Amazon_S3_Demo_All
- ✓ File path Type : Wildcard file path



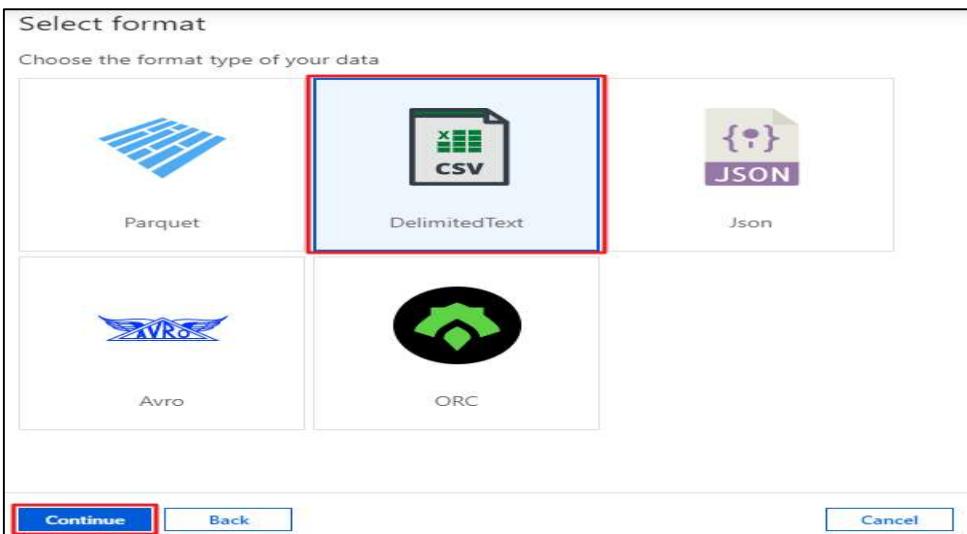
13. Sink탭으로 이동하여 Sink dataset의 +New를 클릭합니다.



14. New Datasource 블레이드가 나타나면 Azure Data Lake Storage Gen2 를 선택 후 Continue 버튼을 클릭합니다.

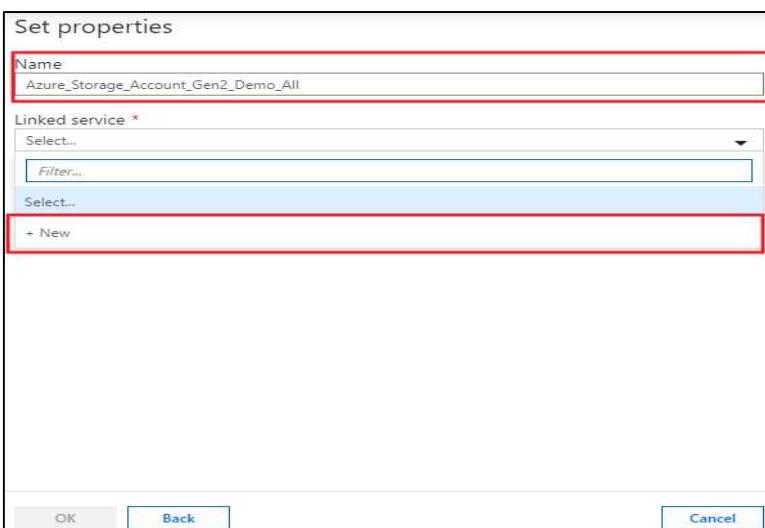


15. format은 DelimitedText를 선택합니다.



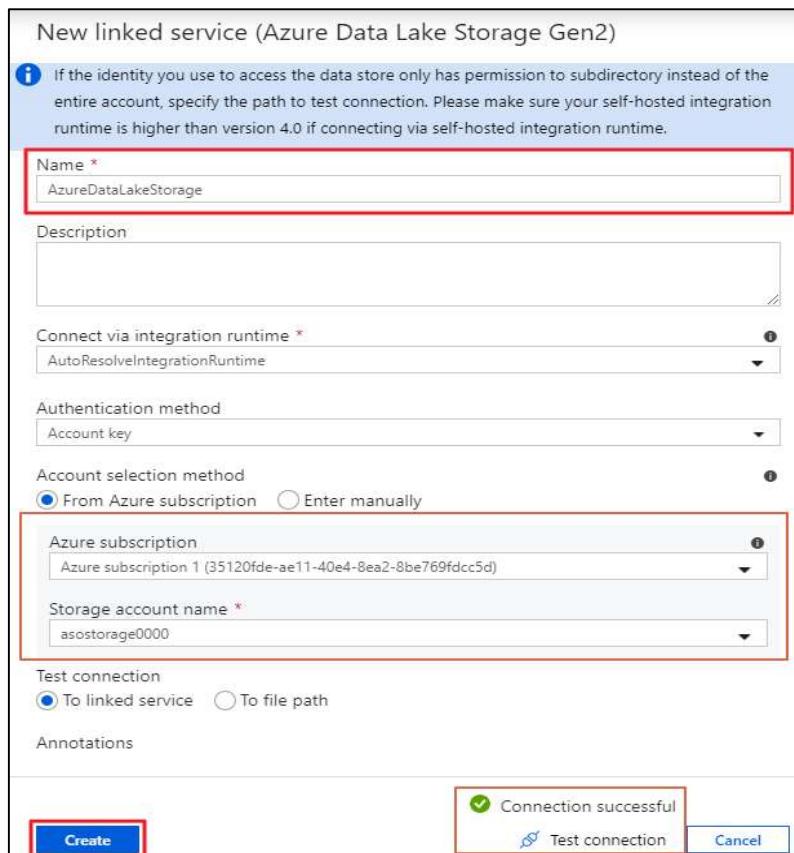
16. Set Properties 블레이드가 열리면 Name 항목을 입력 후 Linked service 의 +New 를 선택합니다.

- ✓ Name: Azure_Storage_Account_Gen2_Demo_All



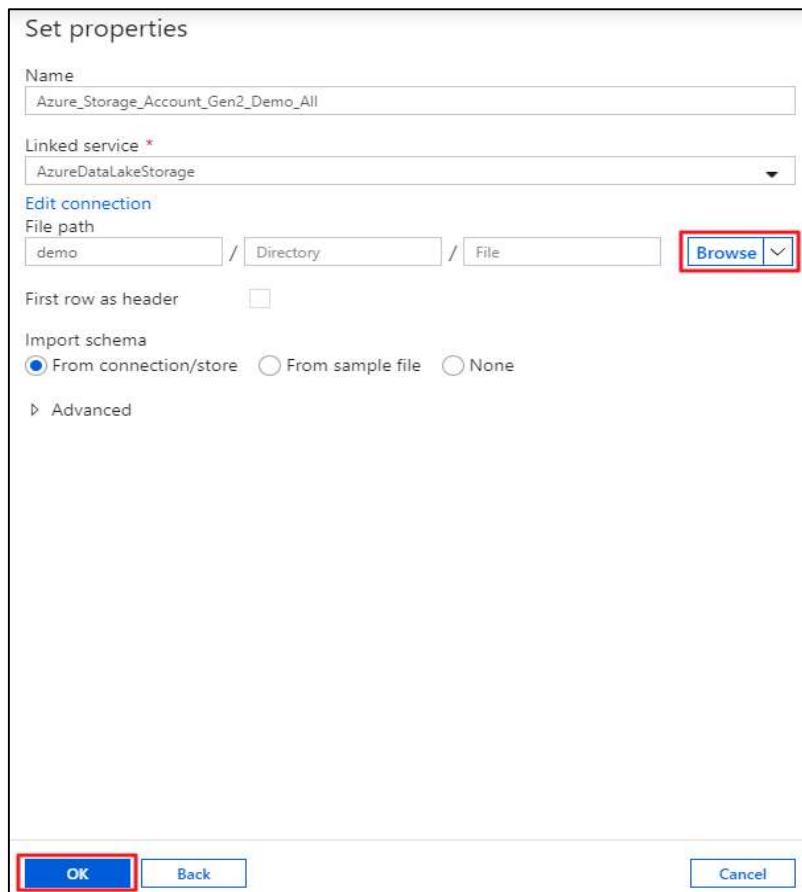
17. New linked service (Azure Data Lake Storage Gen2) 블레이드가 열리면, 아래와 같이 정보를 입력하고 Test Connection 버튼을 클릭하여 Connection successful 메시지가 출력되면 Create 버튼을 클릭합니다.

- ✓ Name : AzureDataLakeStorage
- ✓ Azure subscription: 본인의 구독
- ✓ Storage account name: asostorageXXXX

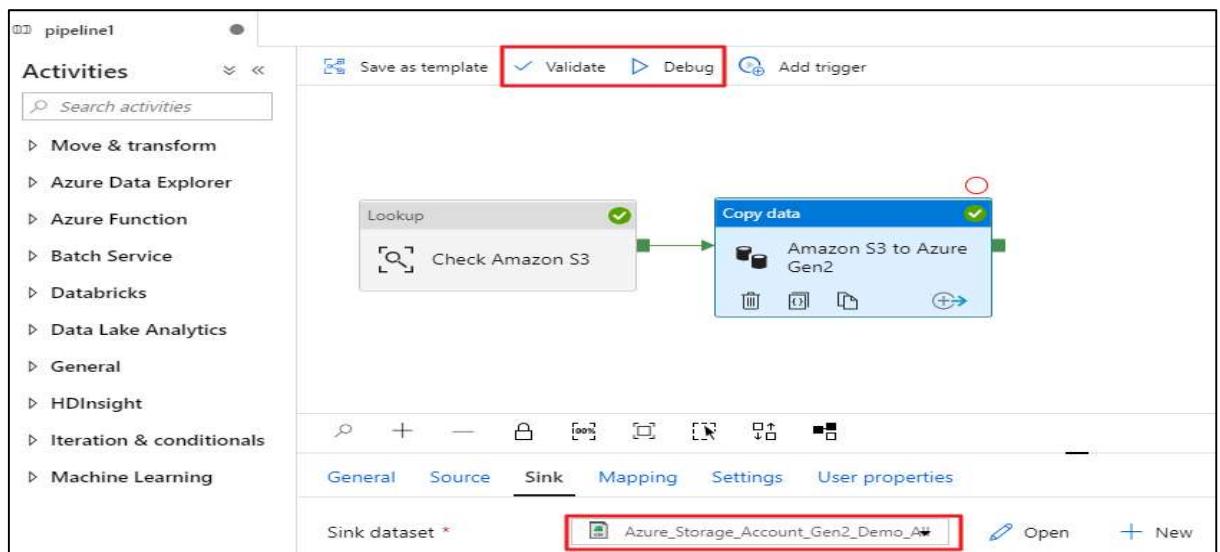


18. Set Properties 블레이드가 열리면 아래의 정보를 입력 후 OK 버튼을 클릭하여 완료합니다.

- ✓ Name : **Azure_Storage_Account_Gen2_Demo_All**
- ✓ Linked service : **AzureDataLakeStorage**
- ✓ File path : **demo /** (browse 버튼 클릭하여 경로 지정)



19. Copy data 액티비티의 Sink탭의 Source Dataset이 세팅된 것을 확인 후 상단의 validate 버튼 클릭 후 Debug 버튼을 클릭하여 실행합니다.



20. 실행된 결과를 확인하기 위해 Azure Portal 의 리소스그룹으로 이동하여 스토리지 계정을 선택합니다.

The screenshot shows the Azure portal interface for the resource group ASO_Edu0000. On the left, there's a navigation pane with options like '개요' (Overview), '활동 로그' (Activity Log), '액세스 제어(IAM)', '태그', '이벤트', '설정' (Settings), '빠른 시작' (Quick Start), '리소스 비용' (Resource costs), '배포' (Deployment), '정책' (Policies), and '속성' (Properties). The main content area displays a list of resources under the heading '리소스 그룹 (변경) : Azure subscription 1'. It includes '구독 ID' (Subscription ID) and '태그 (변경)' (Tags). A search bar at the top right allows filtering by name, type, and location. The list contains several items: 'ASODatafactory0000' (Data Factory(V2)), 'asodwserver0000' (SQL Server), 'asostorage0000' (Storage Account, highlighted with a red border), and 'CohoDW(asodwserver0000/CohoDW)' (SQL Data Warehouse).

21. 컨테이너 버튼을 클릭합니다.

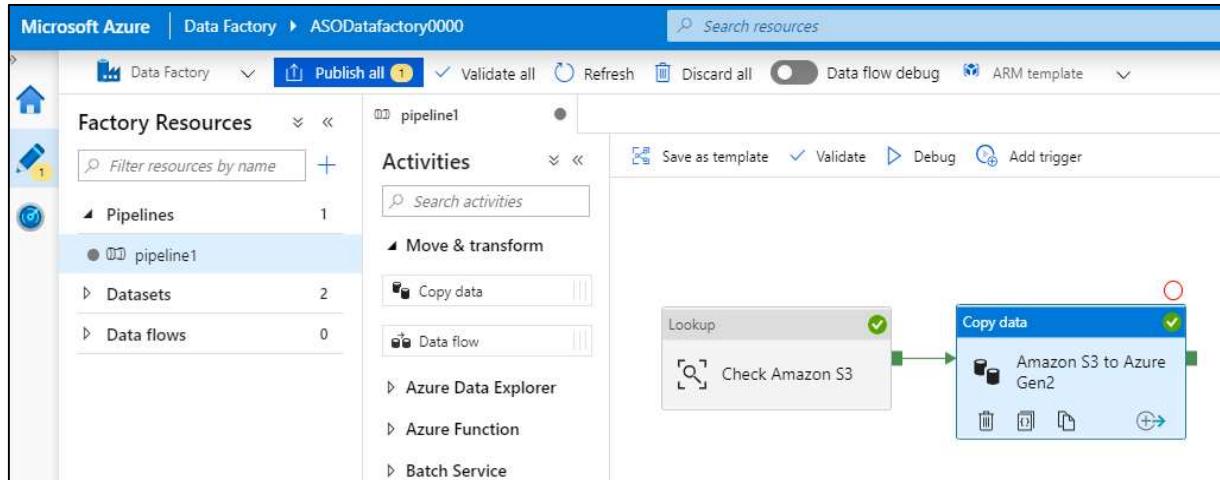
The screenshot shows the Azure portal interface for the storage account asostorage0000. The left sidebar has options like '개요' (Overview), '활동 로그' (Activity Log), '액세스 제어(IAM)', '태그', '문제 진단 및 해결' (Diagnose & solve problems), '데이터 전송' (Data transfer), '이벤트' (Events), 'Storage Explorer(미리 보기)' (Storage Explorer Preview), '설정' (Settings), and '액세스 키' (Access keys). The main content area shows details for the storage account, including its subscription, location, and tags. A large red box highlights the '컨테이너' (Container) section, which describes it as a '대규모로 확장 가능한 데이터 레이크 스토리지' (A highly scalable data lake storage). To the right, there's a '파일 공유' (File sharing) section with a link to '자세한 정보' (Detailed information).

22. 컨테이너 목록에서 demo를 선택하여 아래와 같이 폴더가 생성 되었는지 확인합니다.

The screenshot shows the Azure portal interface for the 'demo' container within the storage account asostorage0000. The left sidebar has options like '개요' (Overview), '액세스 제어(IAM)', '설정' (Settings), and '메타데이터' (Metadata). The main content area shows the '업로드' (Upload) button, '디렉터리 추가' (Add directory), '새로 고침' (Refresh), and '이름 바꾸기' (Rename) buttons. It also displays the authentication method as '액세스 키 (Azure AD 사용자 계정으로 전환)' (Access key (Switch to Azure AD user account)) and the location as 'demo'. A search bar at the top right allows searching for blobs by name or prefix. Below, a table lists blobs: 'mysql' (unstructured), 'structured' (unstructured), and 'unstructured' (unstructured). The 'mysql' blob is highlighted with a red border.

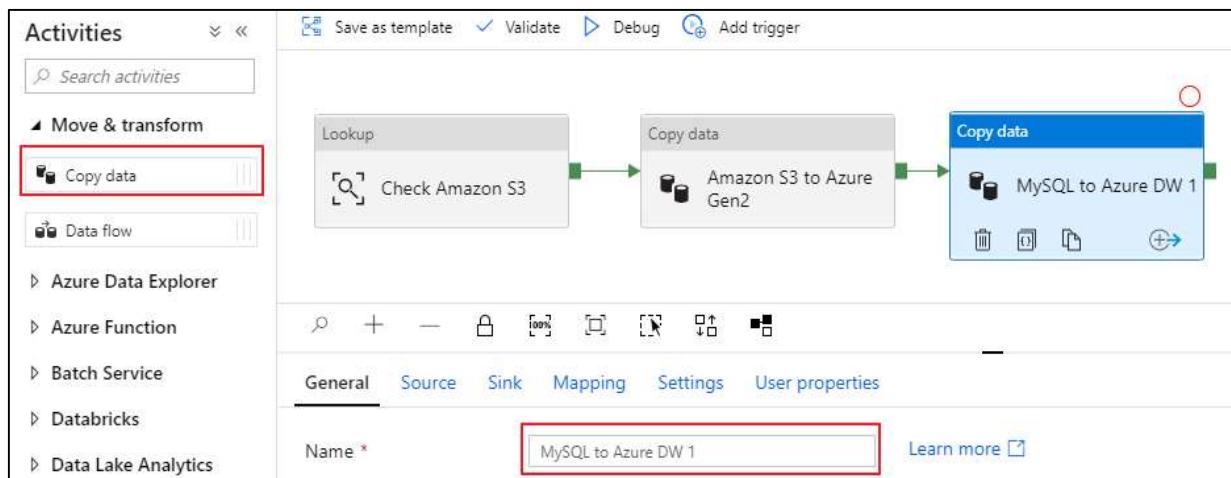
4.2. Task 2: MySQL to Azure Synapse Analytics

1. Data Factory 포털 화면으로 이동합니다.

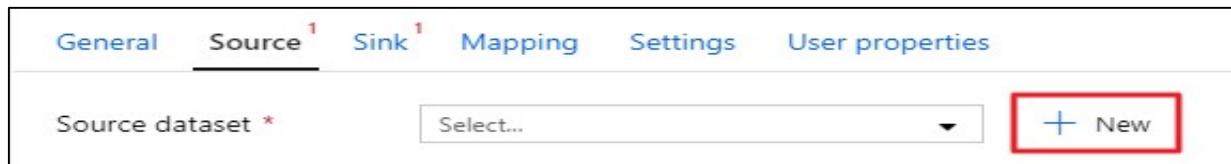


2. Activities > Move & transform > Copy data를 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.

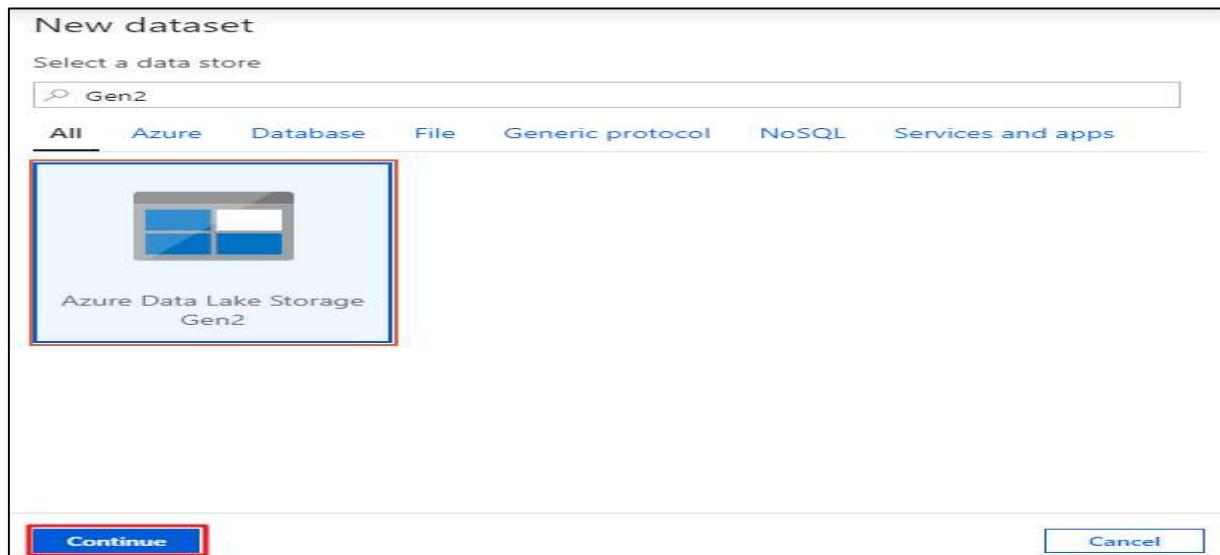
- ✓ Name: MySQL to Azure DW 1



3. Source 탭으로 이동하여 Source dataset의 +New 를 클릭합니다.



4. New dataset 블레이드에서 Gen2를 검색하여 Azure Data Lake Storage Gen2를 선택 후 Continue를 클릭합니다.

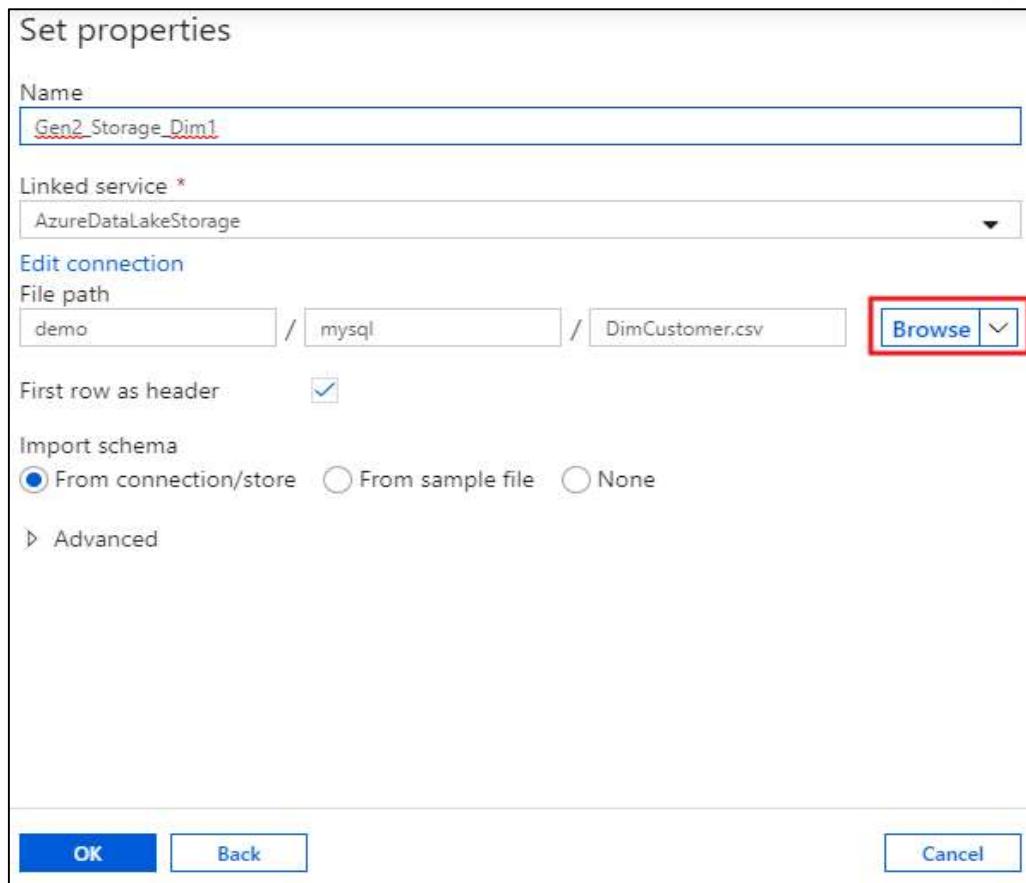


5. format은 DelimitedText를 선택합니다

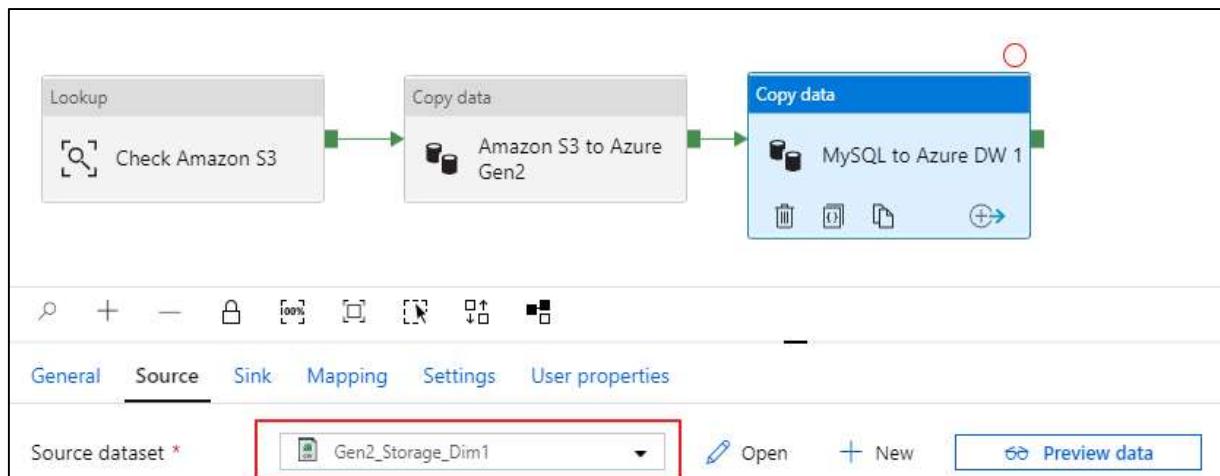


6. Set Properties 블레이드가 열리면 아래의 정보를 입력 후 OK버튼을 클릭합니다.

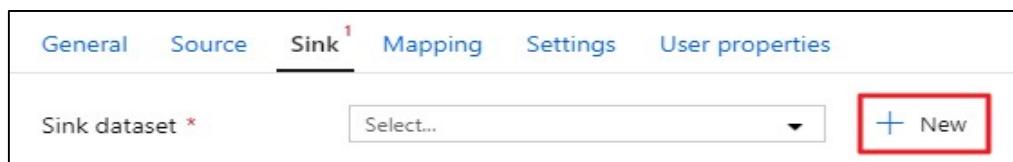
- ✓ Name: **Gen2_Storage_Dim1**
- ✓ Linked service : **AzureDataLakeStorage**
- ✓ File path : **demo / mysql / DimCustomer.csv** (browse 버튼 클릭하여 경로 지정)
- ✓ First row as header : **Check**



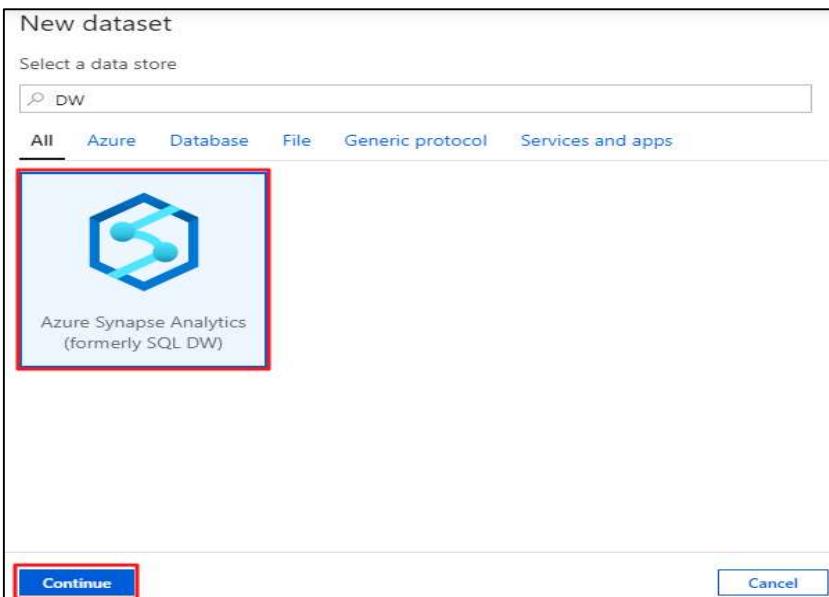
7. Source 탭의 **Source Dataset**이 세팅된 것을 확인합니다.



8. Sink 탭으로 이동하여 Sink dataset의 **+New**를 클릭합니다.

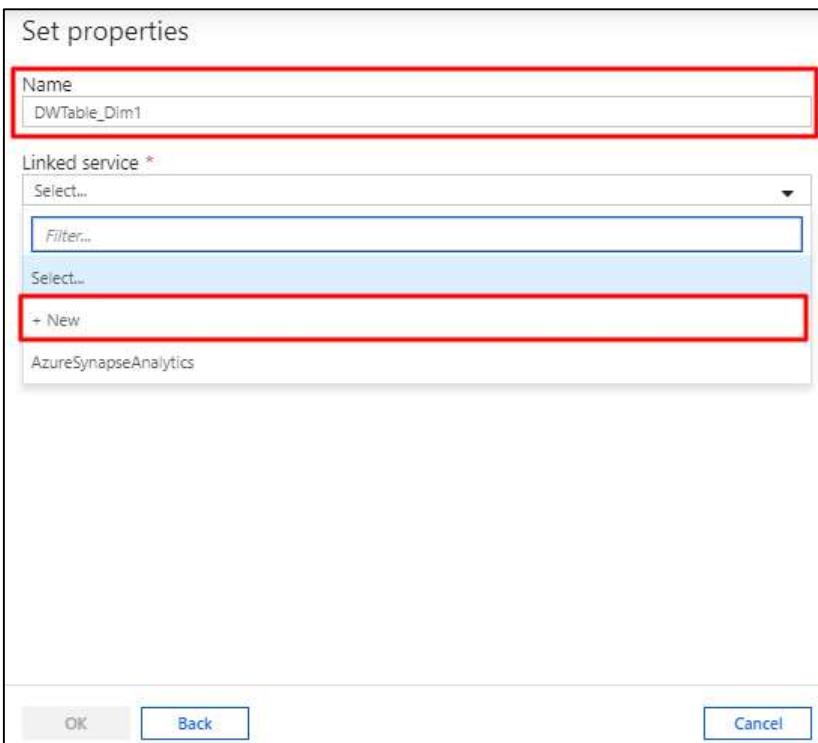


9. New dataset 블레이드에서 DW를 검색하여 **Azure Synapse Analytics**를 선택 후 **Continue**를 클릭합니다.



10. Set Properties 블레이드가 열리면 Name 항목을 입력 후 Linked service 항목을 클릭한 후 +New 를 선택합니다.

- ✓ Name: DWTable_Dim1



11. New linked service (Azure Synapse Analytics (formerly SQL DW)) 블레이드가 열리면, 아래와 같이 정보를 입력하고 Test Connection 버튼을 클릭하여 Connection successful 메시지가 출력되면 Create 버튼을 클릭합니다.

- ✓ Name : AzureSynapseAnalytics
- ✓ Azure subscription : 본인의 구독
- ✓ Server name : asodwserverXXXX (DW 계정 명)

- ✓ Database name : CohoDW
- ✓ User name : demouser
- ✓ Password : Demo@pass123

New linked service (Azure Synapse Analytics (formerly SQL DW))

Name *
AzureSynapseAnalytics

Description

Connect via integration runtime *
AutoResolveIntegrationRuntime

Connection string	Azure Key Vault
-------------------	-----------------

Account selection method
 From Azure subscription Enter manually

Azure subscription
Azure subscription 1 (35120fde-ae11-40e4-8ea2-8be769fdcc5d)

Server name *
asodwserver0000

Database name *
CohoDW

Authentication type *
SQL authentication

User name *
demouser

Password	Azure Key Vault
----------	-----------------

Password *

Create
Connection successful
Test connection
Cancel

12. Set properties 화면이 열리면 아래와 같이 입력 후 OK 버튼을 클릭합니다.

- ✓ Name: DWTable_Dim1
- ✓ Linked service : AzureSynapseAnalytics
- ✓ Table Name : dbo.DimCustomer

Set properties

Name
DWTable_Dim1

Linked service *
AzureSynapseAnalytics

Edit connection

Table name
dbo.DimCustomer

Edit

Import schema
 From connection/store None

► Advanced

OK Back Cancel

13. Copy data 액티비티의 Sink탭의 Source Dataset이 세팅된 것을 확인합니다.

General Source Sink Mapping Settings User properties

Sink dataset * DWTable_Dim1

14. Settings 탭으로 이동하여 Enable staging 항목에 체크 후 Staging account linked service 항목의 +New를 클릭합니다.

General Source Sink Mapping Settings User properties

Info You will be charged # of used DIUs * copy duration * \$0.25/DIU-hour. Local currency and

Data integration unit Auto Edit

Degree of copy parallelism Edit

Fault tolerance Abort activity on first incompatible row Edit

Enable staging

Add dynamic content [Alt+P]

▲ Staging settings

Staging account linked service Select...

15. 다음 내용을 입력 후 **Test connection**을 클릭하여 연결을 확인한 후 **Create** 버튼을 클릭합니다.

- ✓ Name: **AzureBlobStorage(임의입력)**
- ✓ Account selection method: **From Azure subscription**
- ✓ Azure subscription: **본인의 구독**
- ✓ Storage account name: **asodwserverXXXX (DW 계정 명)**

New linked service

If the identity you use to access the data store only has permission to subdirectory instead of the entire account, specify the path to test connection. Please make sure your self-hosted integration runtime is higher than version 4.0 if connecting via self-hosted integration runtime.

Name *
AzureBlobStorage

Description

Connect via integration runtime *
AutoResolveIntegrationRuntime

Authentication method
Account key

Connection string Azure Key Vault

Account selection method
 From Azure subscription Enter manually

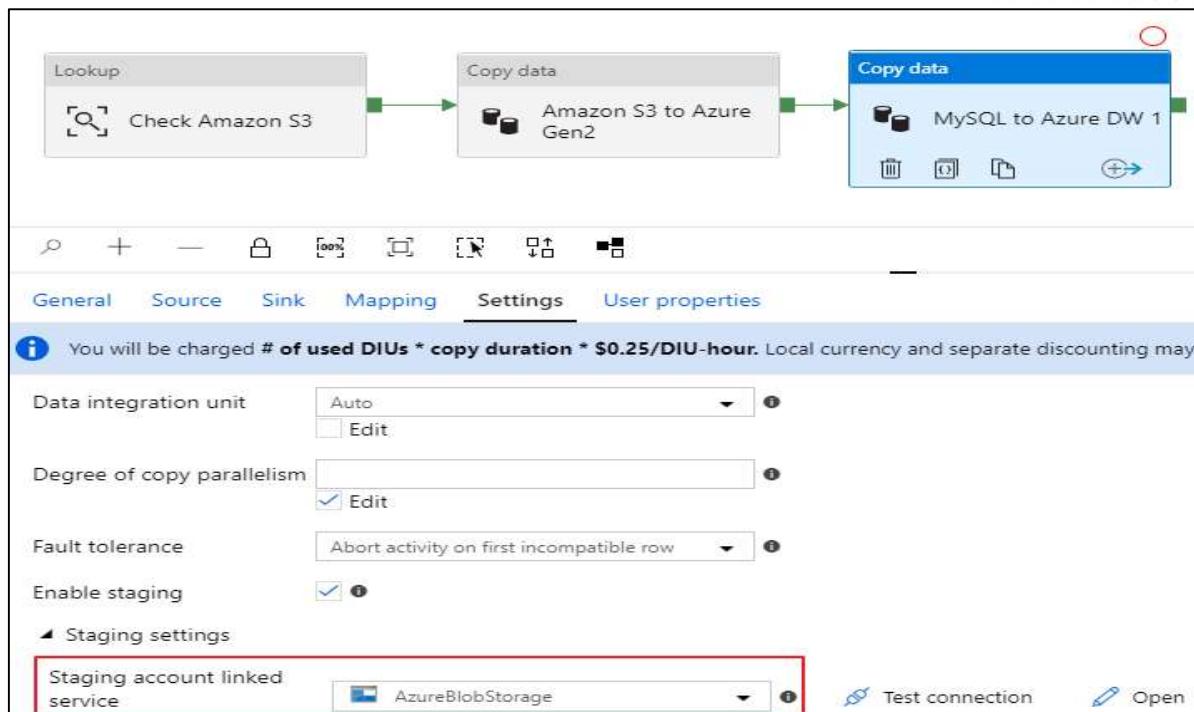
Azure subscription
Azure subscription 1 (35120fde-ae11-40e4-8ea2-8be769fdcc5d)

Storage account name *
asostorage0000

Additional connection properties

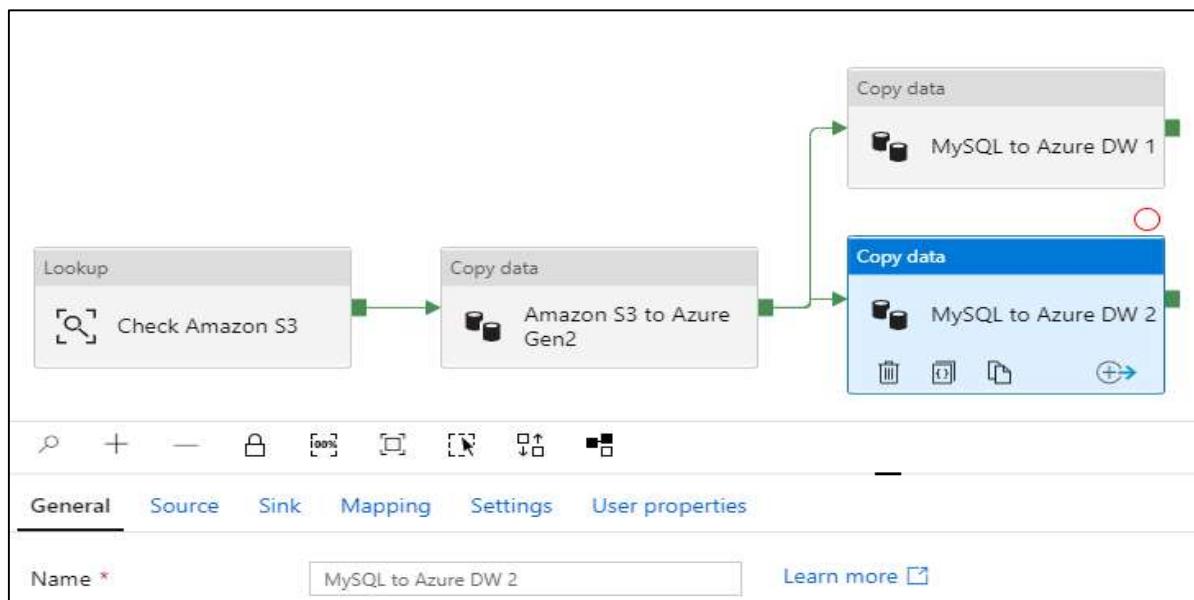
Create **Connection successful** **Test connection** **Cancel**

16. Settings 탭으로 이동하여 Enable staging 항목에 체크 후 Staging account linked service 항목을 **AzureBlobStorage**로 선택합니다.



17. Activities > Move & transform > Copy data를 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.

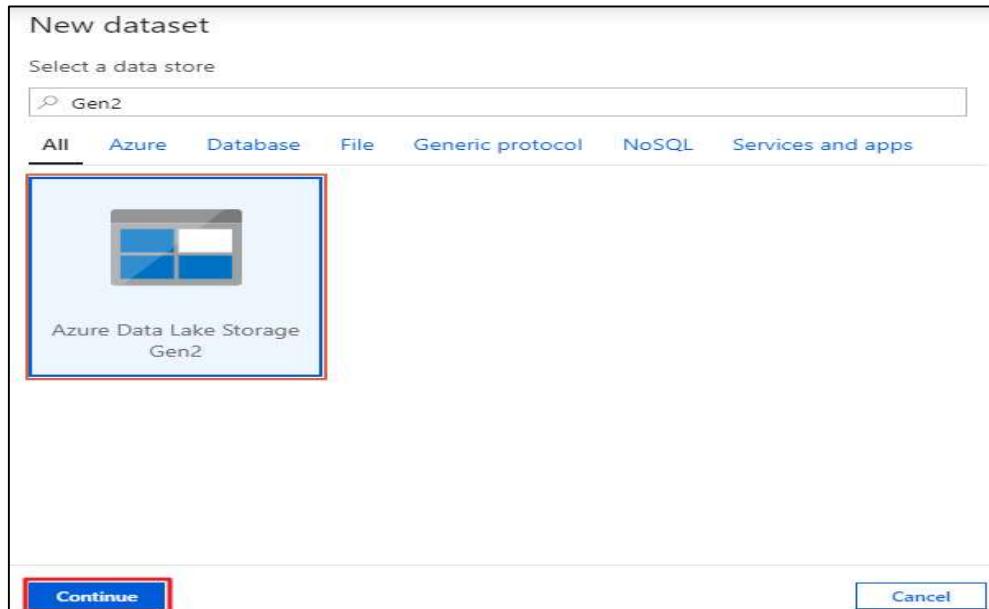
- ✓ Name: MySQL to Azure DW 2



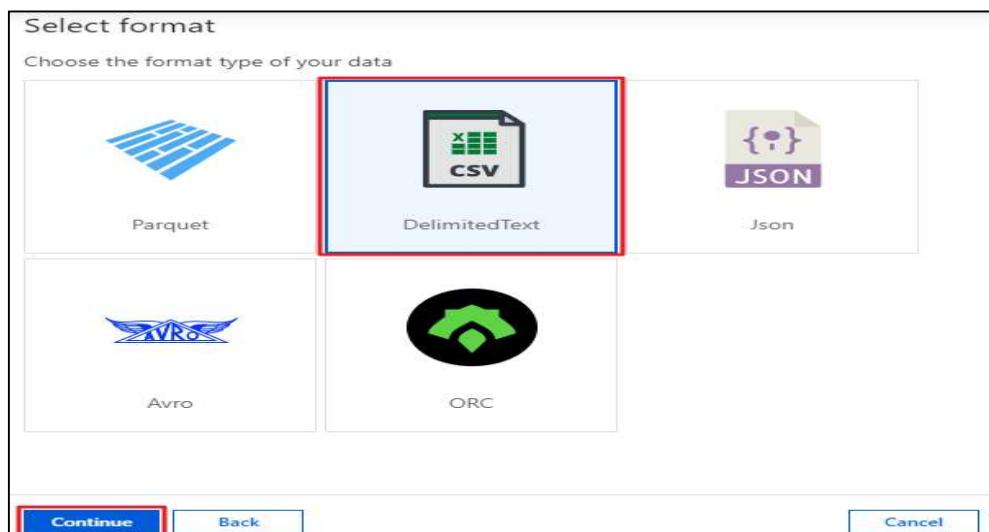
18. Source 탭으로 이동하여 Source dataset의 +New 를 클릭합니다.

General	Source	Sink	Mapping	Settings	User properties
Source dataset *					
<input type="button" value="Select..."/> <input style="border: 2px solid red;" type="button" value="+ New"/>					

19. New dataset 블레이드에서 Gen2를 검색하여 Azure Data Lake Storage Gen2를 선택 후 Continue를 클릭합니다.

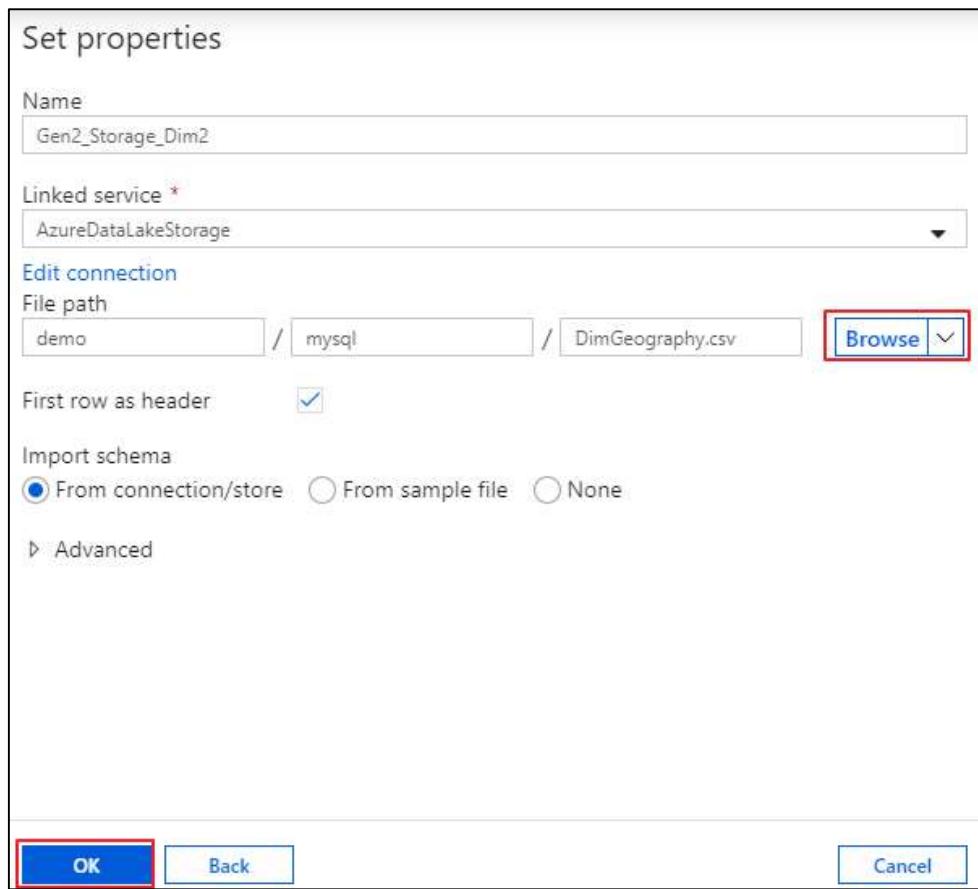


20. format은 DelimitedText를 선택합니다



21. Set Properties 블레이드가 열리면 아래의 정보를 입력 후 OK버튼을 클릭합니다.

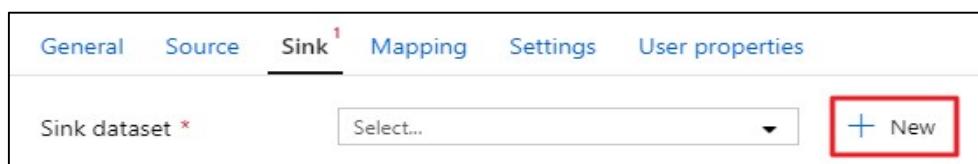
- ✓ Name: **Gen2_Storage_Dim2**
- ✓ Linked service : **AzureDataLakeStorage**
- ✓ File path : **demo / mysql / DimGeography.csv**
- ✓ First row as header : **Check**



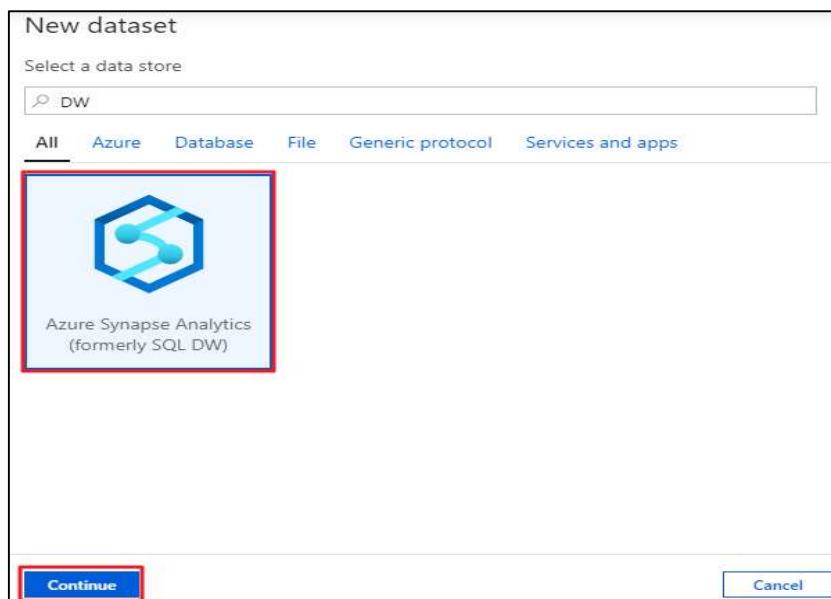
22. Source탭의 **Source Dataset**이 세팅된 것을 확인합니다.



23. Sink탭으로 이동하여 Sink dataset의 **+New**를 클릭합니다.

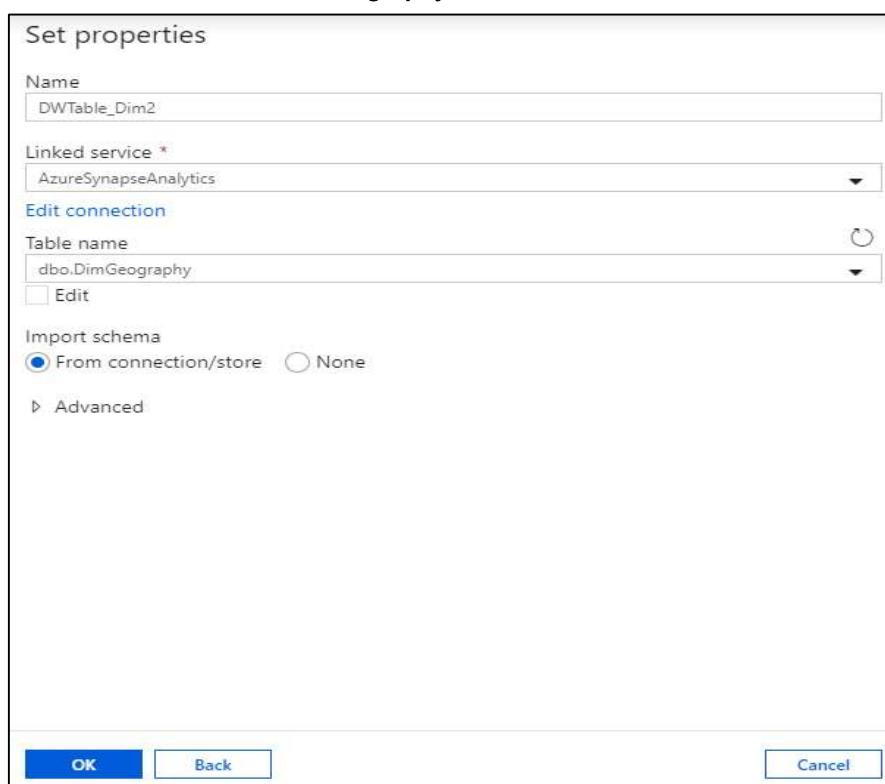


24. New dataset 블레이드에서 DW를 검색하여 **Azure Synapse Analytics**를 선택 후 **Continue**를 클릭합니다.



25. Set properties 화면이 열리면 아래와 같이 입력 후 OK 버튼을 클릭합니다.

- ✓ Name: **DWTable_Dim2**
- ✓ Linked service : **AzureSynapseAnalytics**
- ✓ Table Name : **dbo.DimGeography**

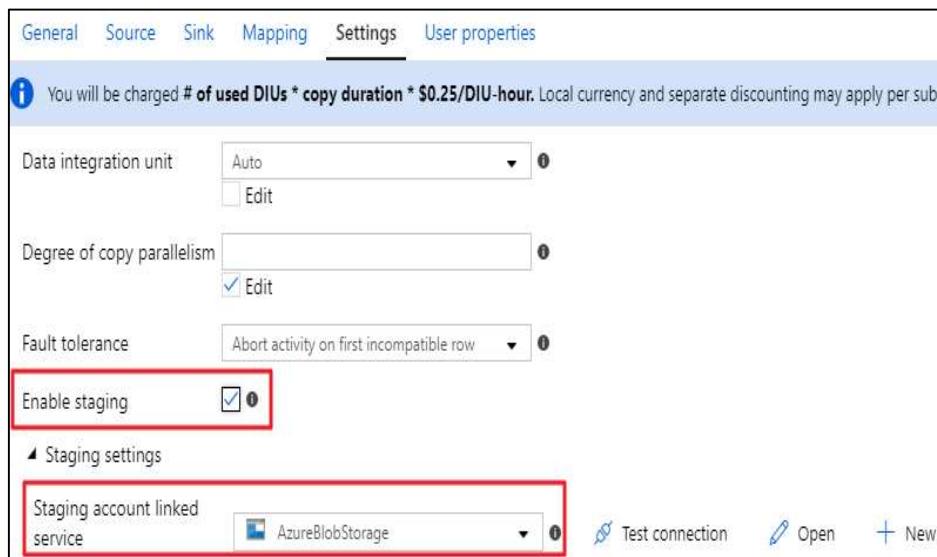


26. Copy data 액티비티의 Sink탭의 Source Dataset이 세팅된 것을 확인합니다.



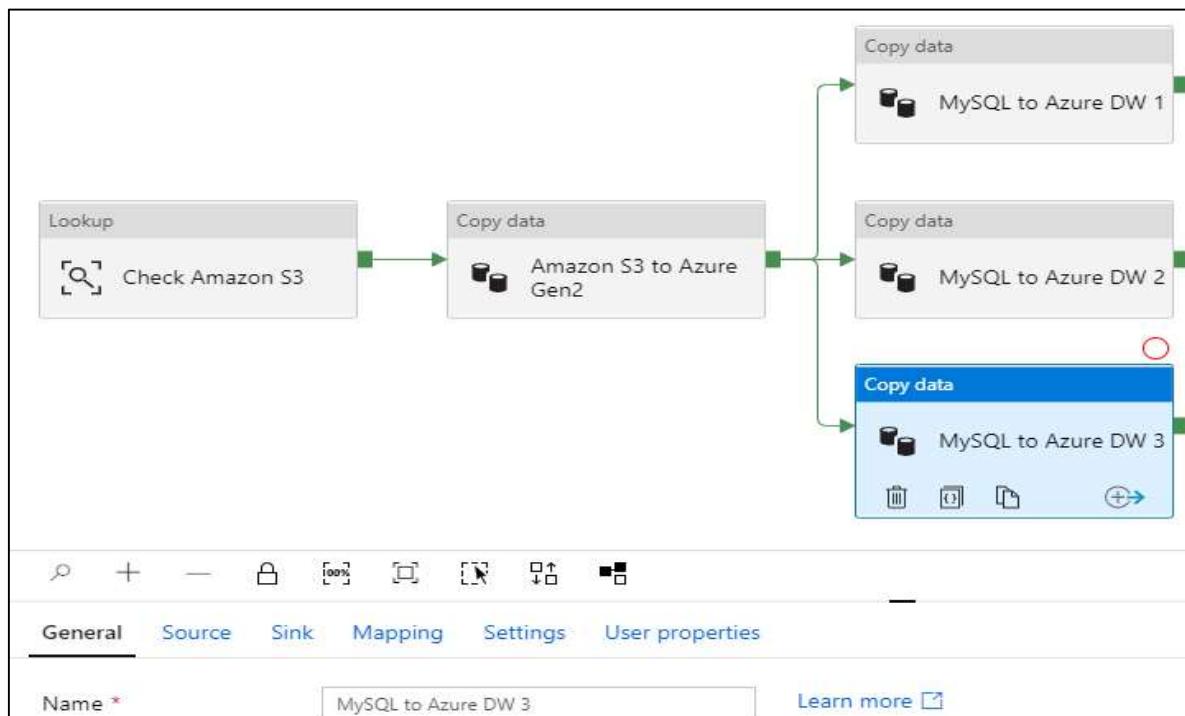
27. Settings 탭으로 이동하여 Enable staging 항목에 체크 후 Staging account linked service 항목은

AzureBlobStorage를 선택합니다.

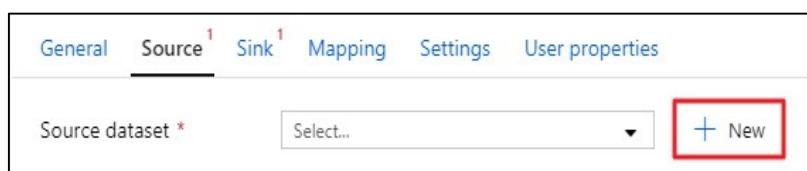


28. Activities > Move & transform > Copy data를 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.

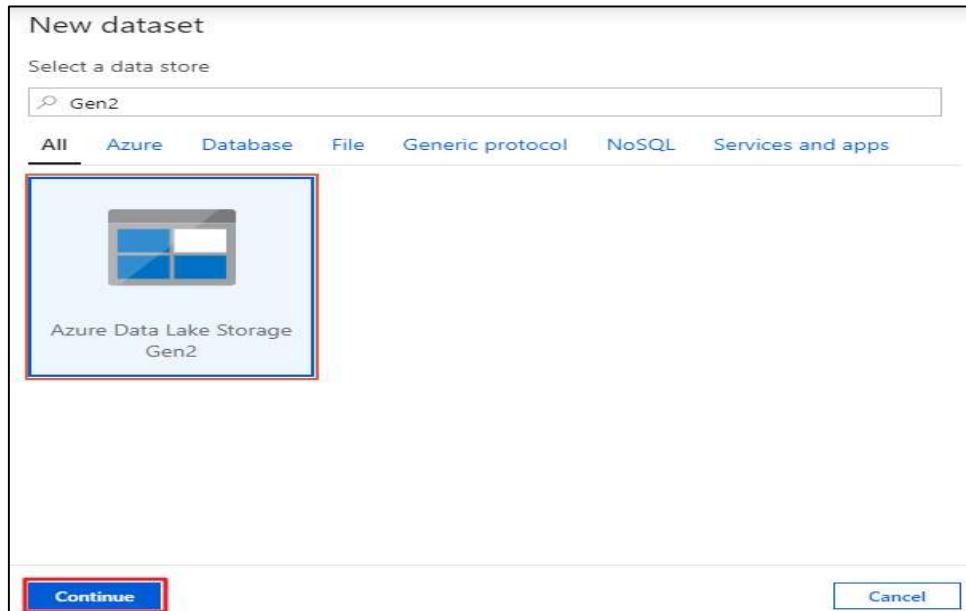
- ✓ Name: MySQL to Azure DW 3



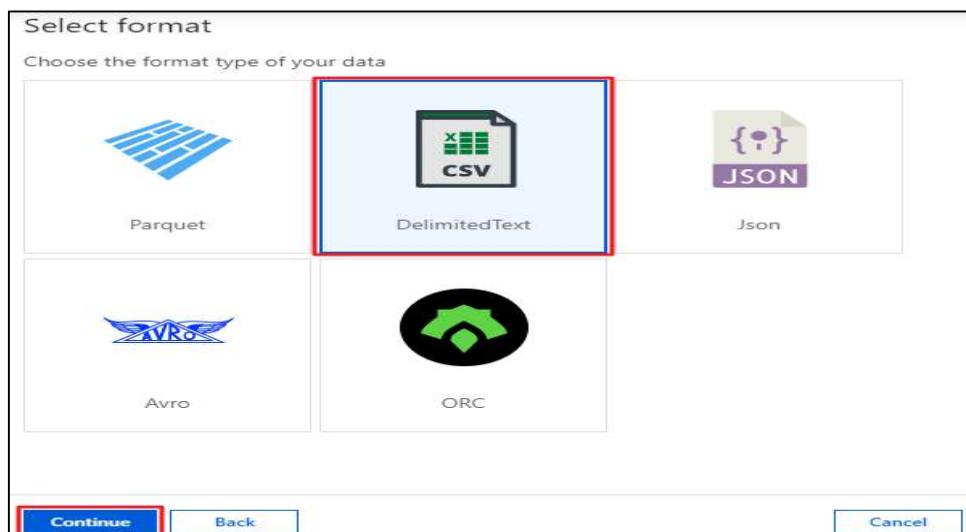
29. Source 탭으로 이동하여 Source dataset의 +New 를 클릭합니다.



30. New dataset 블레이드에서 Gen2를 검색하여 Azure Data Lake Storage Gen2를 선택 후 Continue를 클릭합니다.

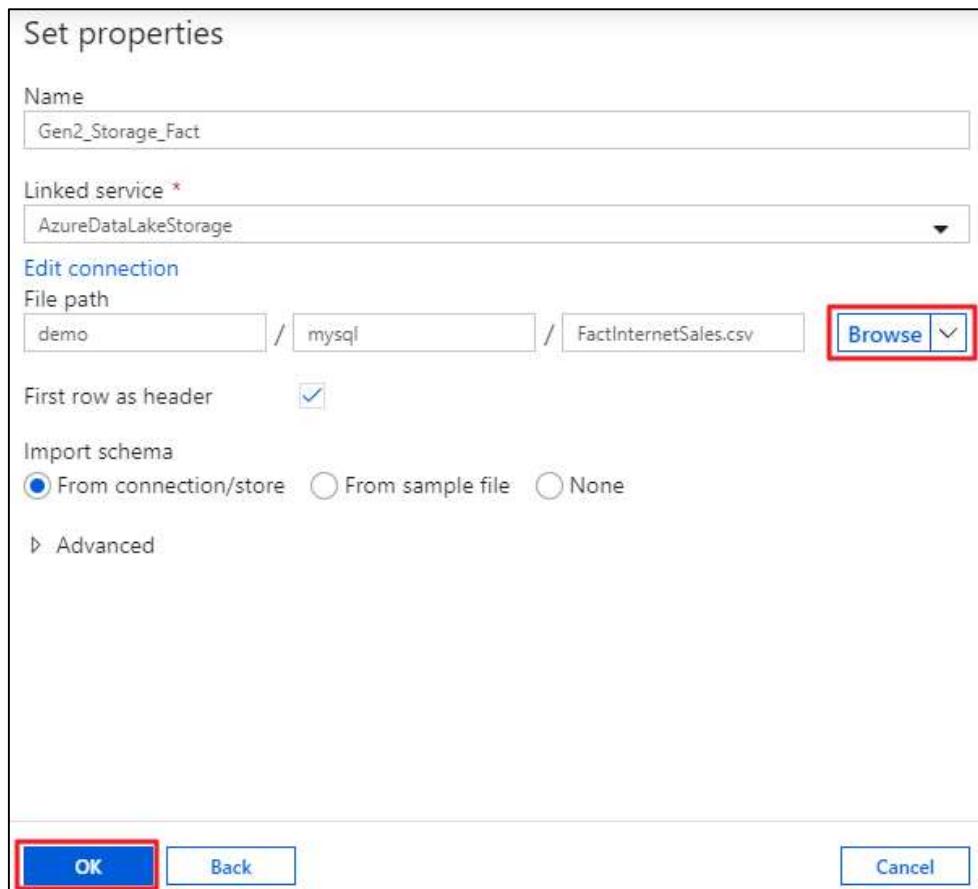


31. format은 DelimitedText를 선택합니다

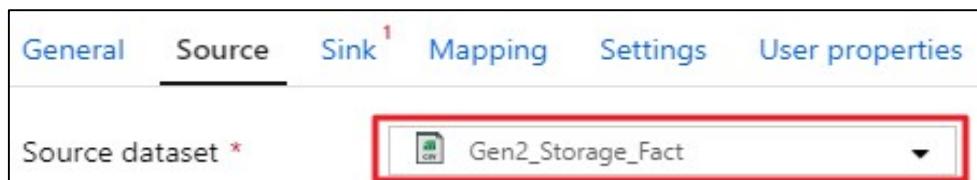


32. Set Properties 블레이드가 열리면 아래의 정보를 입력 후 OK버튼을 클릭합니다.

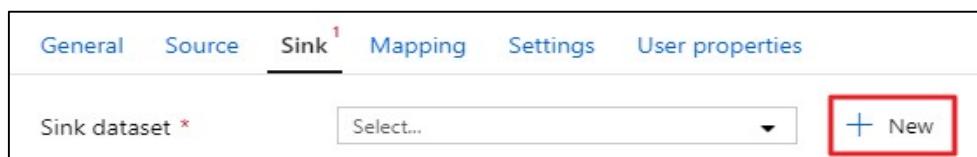
- ✓ Name: **Gen2_Storage_Fact**
- ✓ Linked service : **AzureDataLakeStorage**
- ✓ File path : **demo / mysql / FactInternetSales.csv** (Browse 버튼 클릭하여 경로지정)
- ✓ First row as header : **Check**



33. Source탭의 Source Dataset이 세팅된 것을 확인합니다.

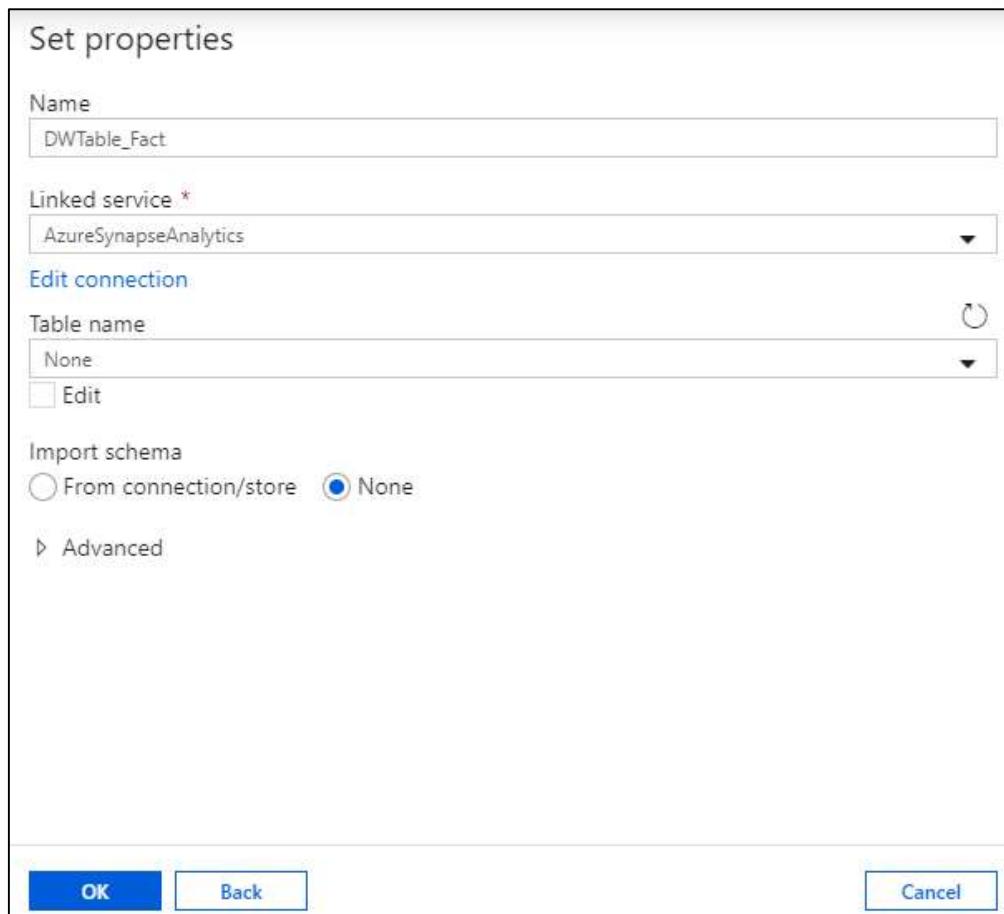


34. Sink탭으로 이동하여 Sink dataset의 +New를 클릭합니다.



35. Set properties 화면이 열리면 아래와 같이 입력 후 OK 버튼을 클릭합니다.

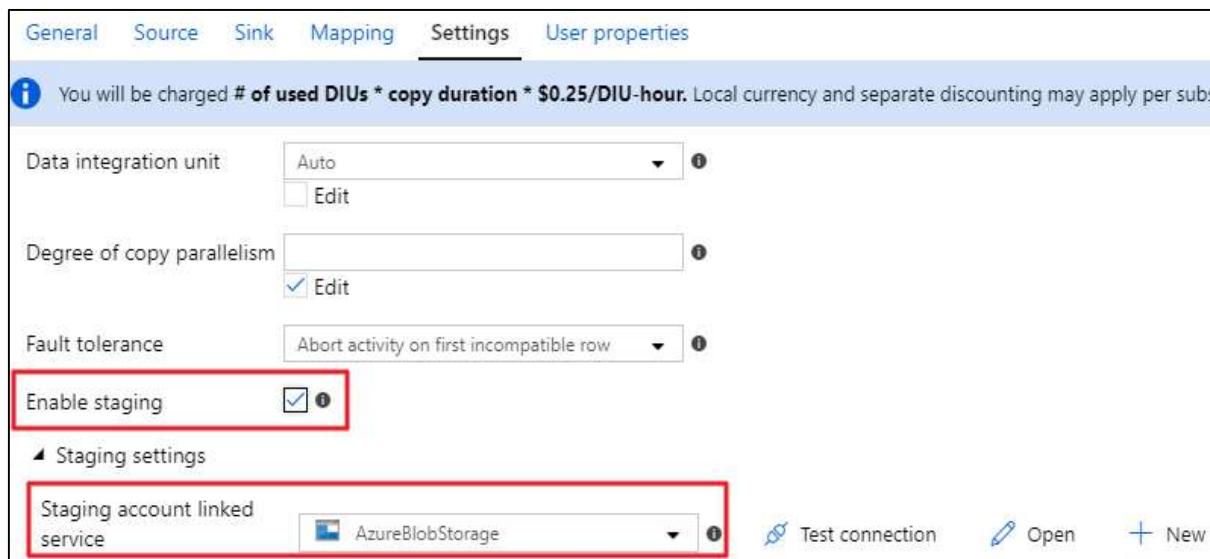
- ✓ Name: DWTable_Fact
- ✓ Linked service : AzureSynapseAnalytics
- ✓ Table Name : dbo.FactInternetsales



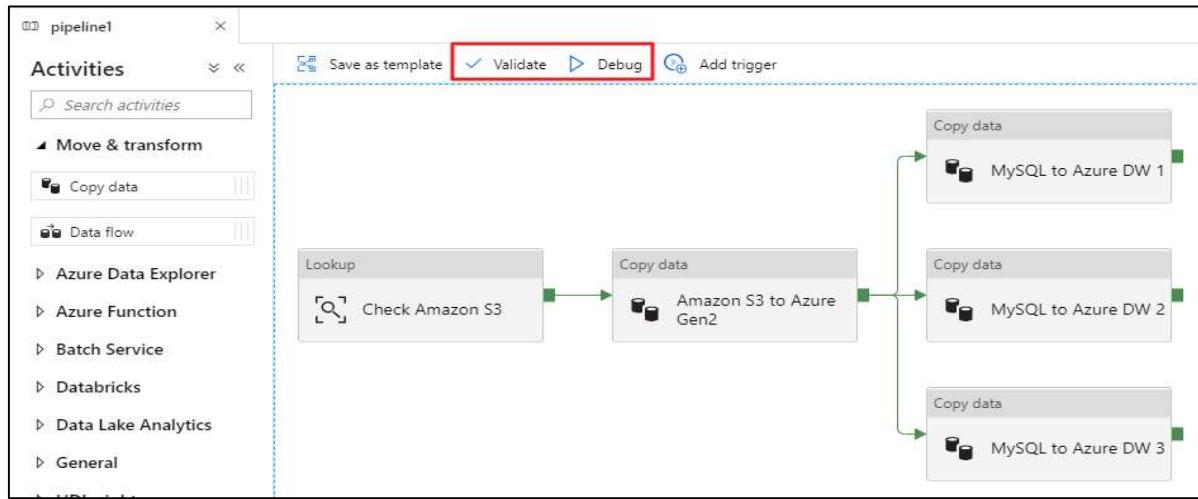
36. Copy data 액티비티의 Sink탭의 Source Dataset이 세팅된 것을 확인합니다.



37. Settings 탭으로 이동하여 Enable staging 항목에 체크 후 Staging account linked service 항목은 AzureBlobStorage를 선택합니다.



38. 상단의 validate 버튼 클릭 후 debug 버튼을 클릭하여 실행합니다.



1. 실행 결과를 확인하기 위해 Azure Portal 의 리소스 그룹으로 이동하여 CohoDW를 선택합니다..

리소스 그룹	구독 (변경) : Azure subscription 1	배포 : 3 성공
개요	구독 ID : 35120fde-ae11-40e4-8ea2-8be769fdcc5d	
활동 토그		
액세스 제어(IAM)		
태그		
이벤트		
설정		
빠른 시작		형식 ↑↓
리소스 비용	ASODatafactory0000	Data Factory(V2)
배포	asodwserver0000	SQL Server
정책	asostorage0000	스토리지 계정
속성	CohoDW(asodwserver0000/CohoDW)	SQL Data Warehouse

2. 왼쪽 메뉴에서 쿼리 편집기를 클릭 후 아래의 정보를 입력하여 로그인합니다.

- ✓ 로그인: demouser
- ✓ 암호: Demo@pass123

3. 다음 쿼리를 1라인씩 블록 지정 하여 실행하여 테이블 별 데이터가 정상적으로 조회되는지 확인합니다.

```
SELECT TOP(100) * FROM [dbo].[DimCustomer]

SELECT TOP(100) * FROM [dbo].[DimGeography]

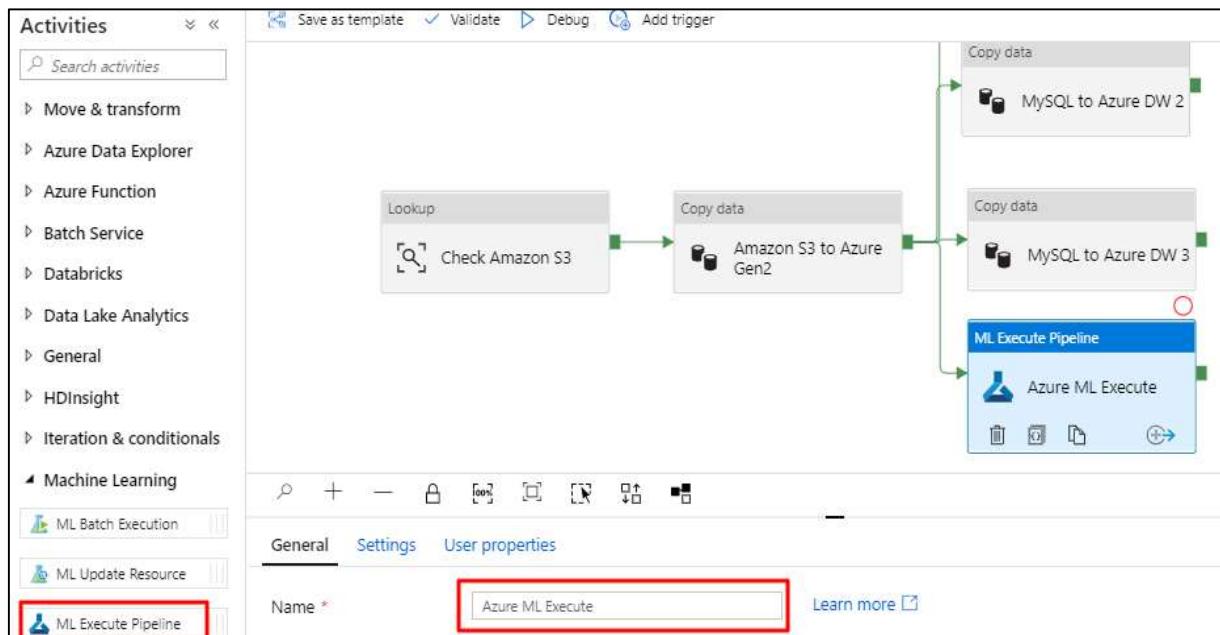
SELECT TOP(100) * FROM [dbo].[FactInternetSales]
```

CustomerKey	GeographyKey	CustomerAlternateKey	Title
14988	248	AW00014988	
15946	298	AW00015946	
15096	156	AW00015096	Mr.
18541	71	AW00018541	
11565	229	AW00011565	
12021	16	AW00012021	

4.3. Task 3: Azure ML Execute

1. Data Factory 포털 화면으로 이동합니다.
2. Activities > Machine Learning > ML Execute Pipeline을 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.

- ✓ Name: Azure ML Execute



3. Setting 탭으로 이동하여 AML Service linked service의 +New를 클릭합니다.



4. New linked service (Azure Synapse Analytics (Azure ML Service)) 블레이드가 열리면, 아래와 같이 정보를 입력하고 Test Connection 버튼을 클릭하여 Connection successful 메시지가 출력되면 Create 버튼을 클릭합니다.

- ✓ Name : AzureMLService
- ✓ Azure ML Service workspace selection method: From Azure subscription
- ✓ Azure subscription : 본인의 구독
- ✓ Azure ML Service workspace name: ASO_MachineLearning
- ✓ Tenant : 56480cdf-0e19-4439-bf83-359f8a87e112
- ✓ Service principal ID: 13c2c33f-271e-4ee0-ad2a-9cf4f6d7b917
- ✓ Service principal key : 40932bbf-4739-4c21-95a1-aa08f1a3e5bc

New linked service (Azure ML Service)

Name *
AzureMLService

Description

Connect via integration runtime *
AutoResolveIntegrationRuntime

Azure ML Service workspace selection method
 From Azure subscription Enter manually

Azure subscription
Azure subscription 1 (35120fde-ae11-40e4-8ea2-8be769fdcc5d)

Azure ML Service workspace name *
ASO_MachineLearning

Tenant *
56480cdf-0e19-4439-bf83-359f8a87e112

Service principal ID *
13c2c33f-271e-4ee0-ad2a-9cf4f6d7b917

Service principal key	Azure Key Vault
Service principal key *	<input type="text"/>

Annotations
+ New

Advanced ●

✓ Connection successful

Create Test connection Cancel

5. AML Service linked service 항목이 세팅된 것을 확인 후 아래의 정보를 입력합니다.

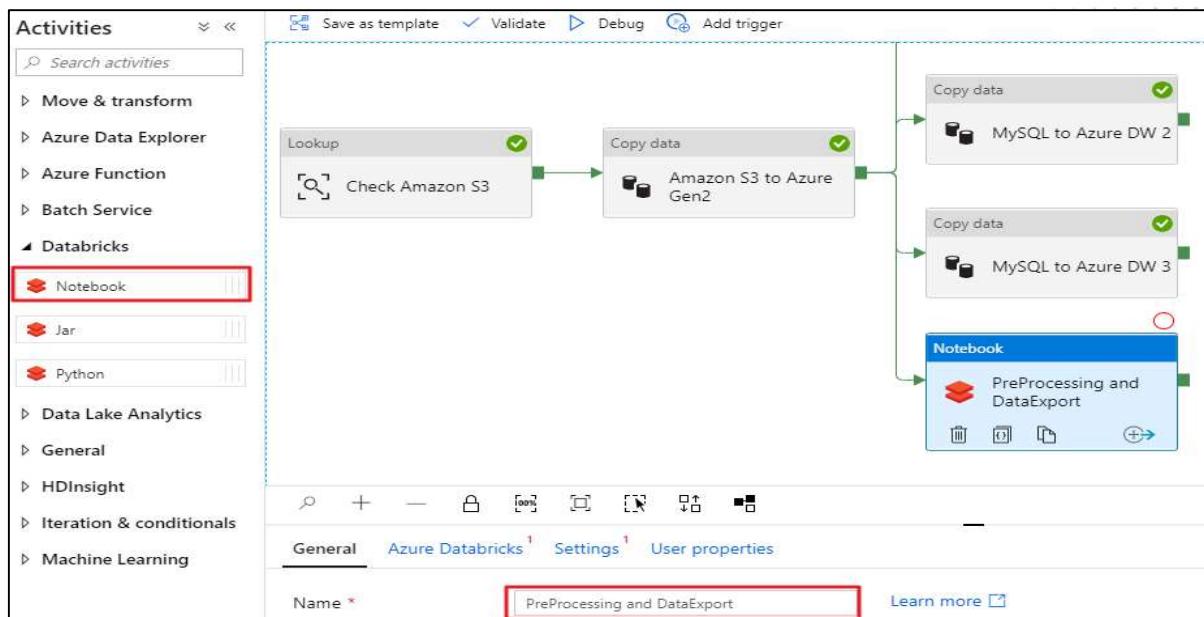
- ✓ ML pipeline name: **Logistic_Movie_Review_Sample**
- ✓ ML pipeline ID: **cd916d3c-6de0-4e90-93a8-b8491d1bd978..**
- ✓ Experiment name : **Designer_test_v1**

General	Settings	User properties
AML Service linked service *	<input type="button" value="AzureMLService"/> ●	Test connection Open + New
ML pipeline name	<input type="button" value="Logistic_Movie_Review_Sample 02-19-2..."/> ●	
ML pipeline ID *	<input type="button" value="cd916d3c-6de0-4e90-93a8-b8491d1bd..."/> ●	Open in Azure Portal
Experiment name	<input type="button" value="Designer_test_v1"/> ●	

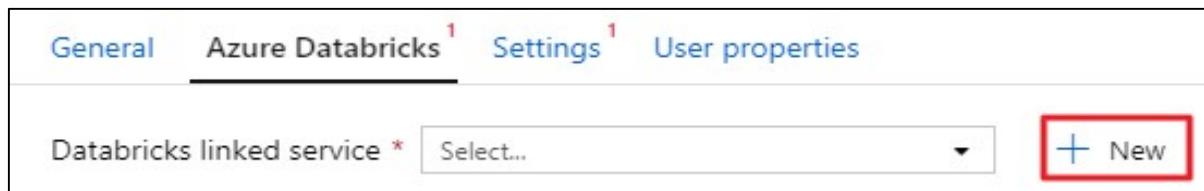
4.4. Task 4: PreProcessing and DataExport

6. Activities > Databricks > Notebook을 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.

- ✓ Name: PreProcessing and DataExport



7. Azure Databricks탭으로 이동하여 Databricks linked service의 +New를 클릭합니다.



8. New linked service (Azure Synapse Analytics (Azure Databricks)) 블레이드가 열리면, 아래와 같이 정보를 입력하고 Test Connection 버튼을 클릭하여 Connection successful 메시지가 출력되면 Create 버튼을 클릭합니다.

- ✓ Name : AzureDatabricks
- ✓ Account selection method: From Azure subscription
- ✓ Azure subscription : 본인의 구독
- ✓ Databricks workspace: ds_Databricks
- ✓ Access token : `dapid1fb7b77a6d0e7f9fb2ccf337367a9a5`
- ✓ Select cluster : Existing interactive cluster
- ✓ Existing cluster ID: Cluster_ASO

New linked service (Azure Databricks)

Name *	AzureDatabricks										
Description											
Connect via integration runtime *	AutoResolveIntegrationRuntime										
Account selection method *	From Azure subscription										
Azure subscription *	Azure subscription 1 (35120fde-ae11-40e4-8ea2-8be769fdcc5d)										
Databricks workspace *	ds_Databricks										
Select cluster	<input type="radio"/> New job cluster <input checked="" type="radio"/> Existing interactive cluster <input type="radio"/> Existing instance pool										
Domain/Region *	https://koreacentral.azure.databricks.net										
<table border="1"> <tr> <td style="background-color: #0072bc; color: white;">Access token</td> <td>Azure Key Vault</td> </tr> <tr> <td>Access token *</td> <td></td> </tr> <tr> <td>.....</td> <td></td> </tr> <tr> <td colspan="2">Choose from existing clusters *</td> </tr> <tr> <td colspan="2">Cluster_ASO</td> </tr> </table>		Access token	Azure Key Vault	Access token *			Choose from existing clusters *		Cluster_ASO	
Access token	Azure Key Vault										
Access token *											
.....											
Choose from existing clusters *											
Cluster_ASO											
<p>Annotations:</p> <p>✓ Connection successful</p> <p>🔗 Test connection Cancel</p>											

9. Databricks linked service 항목이 세팅된 것을 확인합니다.

General	Azure Databricks	Settings ¹	User properties
<p>Databricks linked service *</p> <p>AzureDatabricks</p>			

10. Settings 탭으로 이동하여 Notebook path 항목의 Browse 버튼을 클릭하여 다음과 같이 경로를 지정합니다.

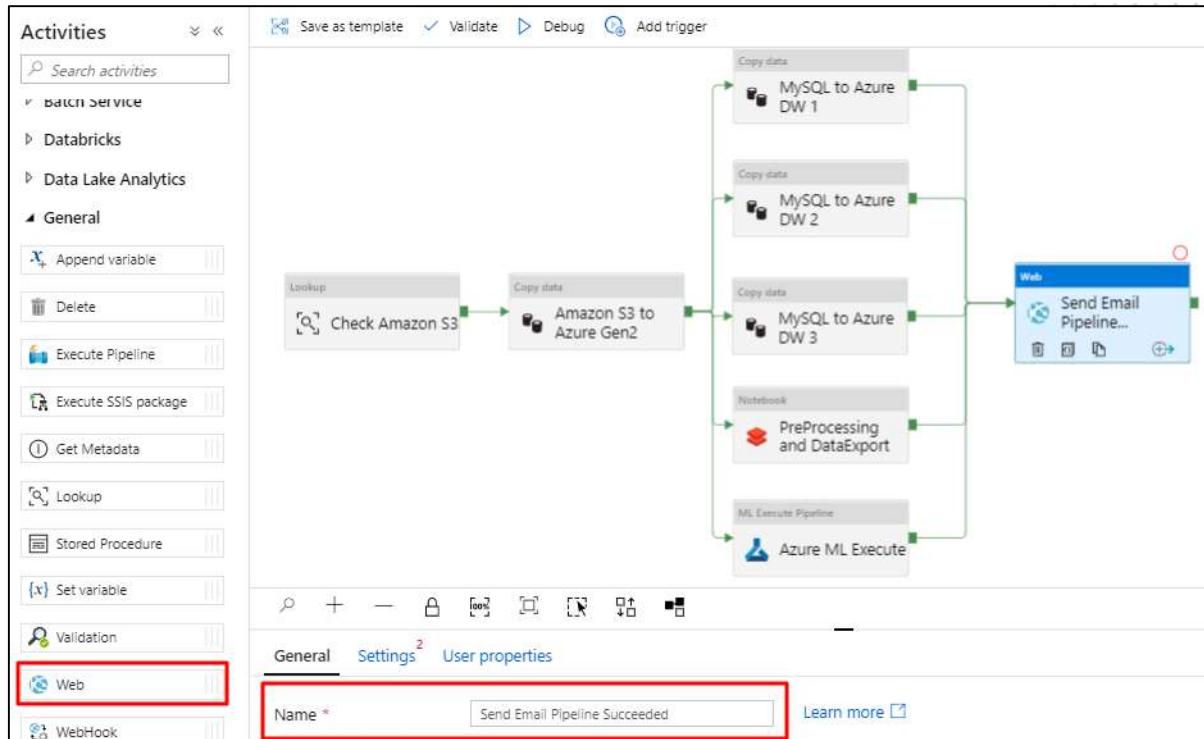
- ✓ Notebook path: /Users/kch8306@datasolutionmlg.onmicrosoft.com/PreProcessing_ASO

General	Azure Databricks	Settings	User properties
Notebook path *	/Users/kch8306@datasolutionmlg.onmicrosoft	Browse	Open
<p>▶ Base parameters</p> <p>▶ Append libraries</p>			

4.5. Task 5: Send Email

1. Activities > General > Web을 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.

- ✓ Name: Send Email Pipeline Succeeded



2. Settings 탭으로 이동하여 아래의 정보를 입력합니다.

- ✓ URL: https://prod-06.koreacentral.logic.azure.com:443/workflows/21f3052675d24be0aa4298a9b0490ce8/triggers/manual/paths/invoke?api-version=2016-10-01&sp=%2Ftriggers%2Fmanual%2Frun&sv=1.0&sig=BfXuXTp05p20wnvDFzkUgi3xbf3XGSB-1qapyC_EDFA
- ✓ Method: POST
- ✓ Headers: +New 클릭 후 Content-Type application/json 추가
- ✓ Body:

```
{"DataFactoryName": "@{pipeline().DataFactory}", "PipelineName": "@{pipeline().Pipeline}", "Subject": "Pipeline Succeeded", "Message": "Succeeded!", "EmailTo": "본인의 메일계정"}
```

General Settings User properties

URL *: https://prod-06.koreacentral.logic.azure.com:443

Method *: POST

Headers *:

NAME	VALUE
Content-Type	application/json

Add dynamic content [Alt+P]

Body:

```
{"DataFactoryName": "@{pipeline().DataFactory}", "PipelineName": "@{pipeline().PipelineName}", "Subject": "Pipeline"}
```

3. 최종 실행을 위해 왼쪽 상단의 **Publish All** 버튼 클릭 후 **Add trigger** 의 Trigger now를 클릭합니다.

Microsoft Azure Data Factory > ASODatafactory0000

Publish all (1) Validate all Refresh Discard all Data flow debug ARM template

Factory Resources Activities

- Pipelines: pipeline1 (selected)
- Datasets: 8
- Data flows: 0

Activities: Pipeline

- Save as template Validate Debug Add trigger (highlighted in red)
- Lookup Trigger now (highlighted in red)
- Move & transform Azure Data Explorer Azure Function

4. 왼쪽 메뉴의 Monitor 아이콘을 클릭하여 진행 상태를 확인합니다.

Microsoft Azure Data Factory > ASODatafactory0000

Moniter

Pipeline runs

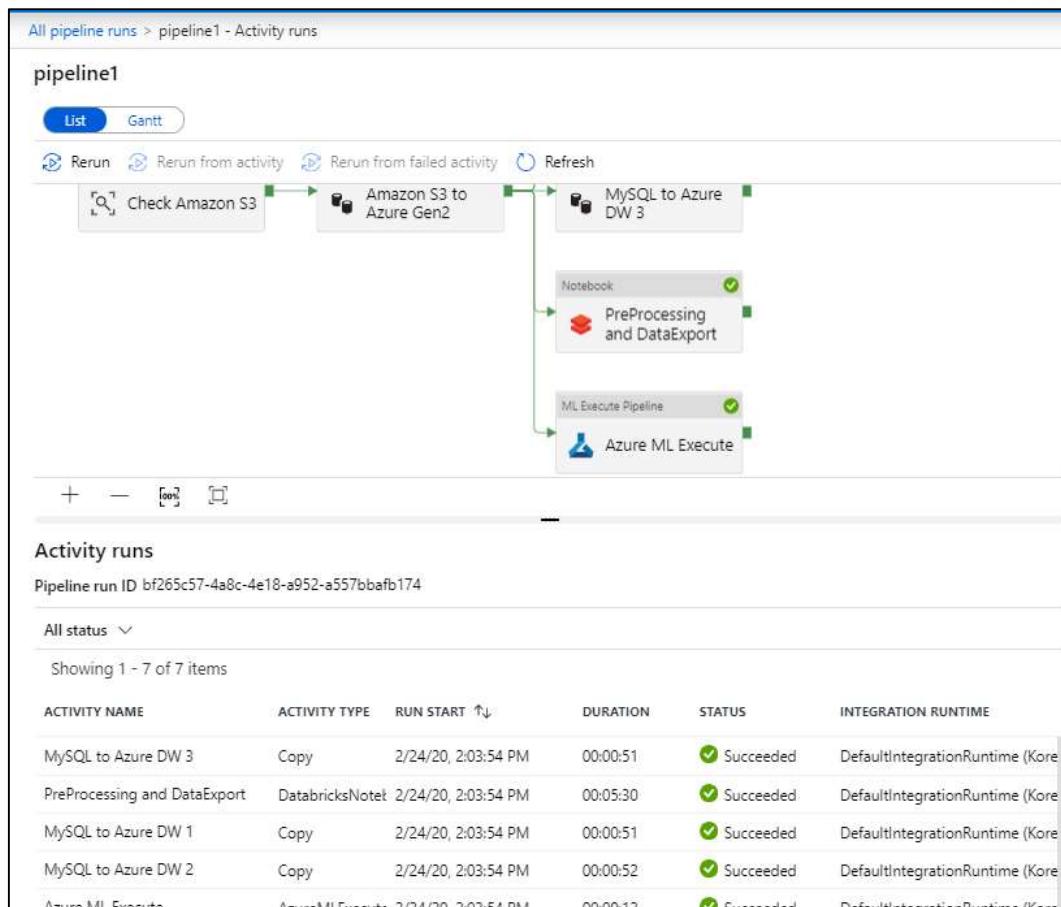
Time : Last 24 hours (2/23/2020 5:04 PM ...) Time zone : Seoul (UTC+9) Runs : Latest runs

All status Rerun Cancel Refresh Edit columns

Showing 1 - 2 items

PIPELINE NAME	RUN START	DURATION	TRIGGERED BY	STATUS
pipeline1	2/24/2020, 2:03:37 PM	00:09:29	Manual trigger	Succeeded

5. PIPELINE NAME을 클릭하여 상세화면으로 이동합니다.

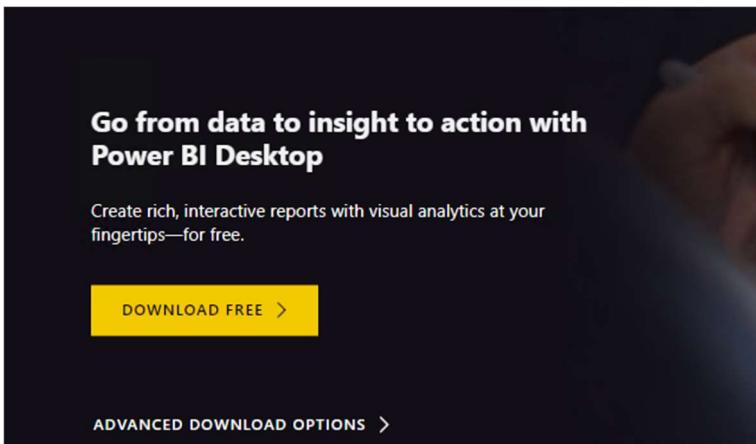


5. Exercise 3: Visualize data with Power BI Desktop

이 실습에서는 Data Warehouse의 데이터를 이용하여 Power BI Desktop과의 통합을 설정합니다.

5.1. Task 1: Install Power BI Desktop

1. 웹 브라우저에서 Power BI Desktop 다운로드 페이지 (<https://powerbi.microsoft.com/en-us/desktop/>)로 이동합니다.
2. 페이지 중간에서 Download Free 링크를 선택합니다.



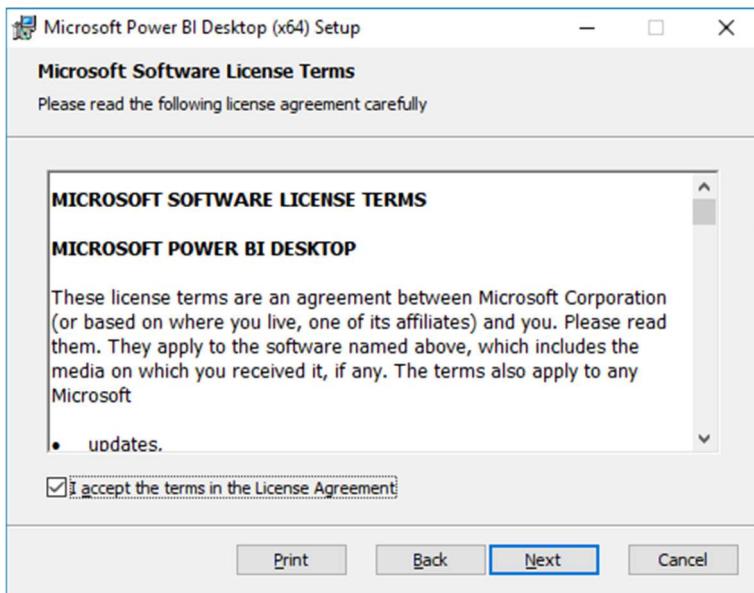
3. 설치파일을 실행합니다.



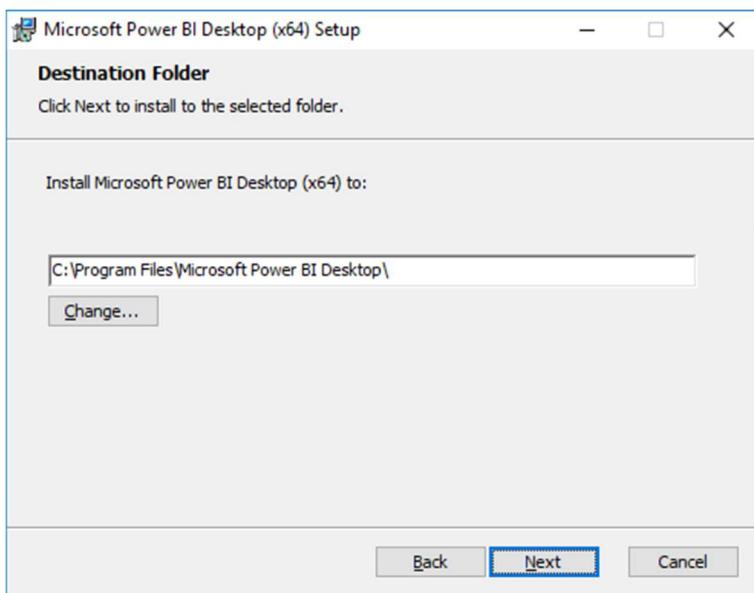
4. Welcome 화면에서 다음을 클릭합니다.



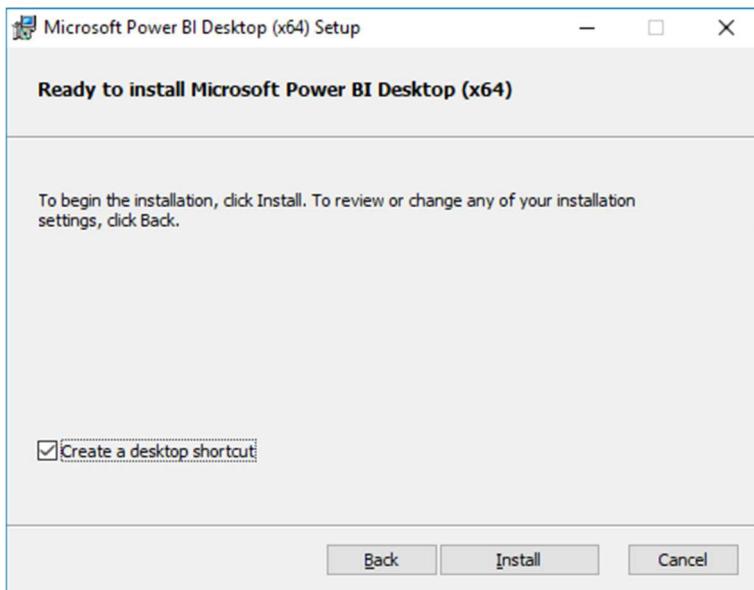
5. 사용권 계약에 동의하고 다음을 선택합니다.



6. 지정 폴더를 디폴트로 두고 다음을 선택합니다.



7. 바탕 화면 바로 가기 만들기 상자가 선택되어 있는지 확인하고 설치를 선택합니다.

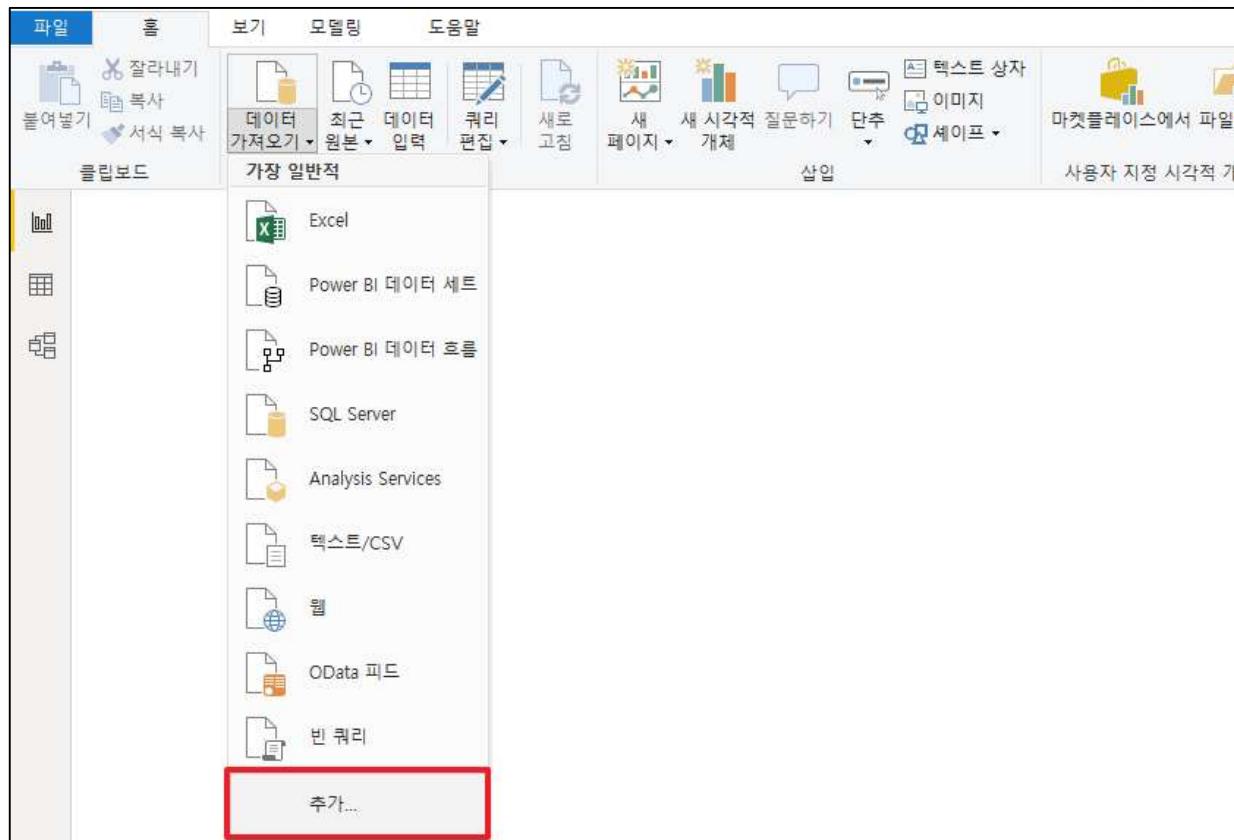


8. Microsoft Power BI Desktop 시작이 선택되어 있는지 확인하고 마침을 선택합니다.

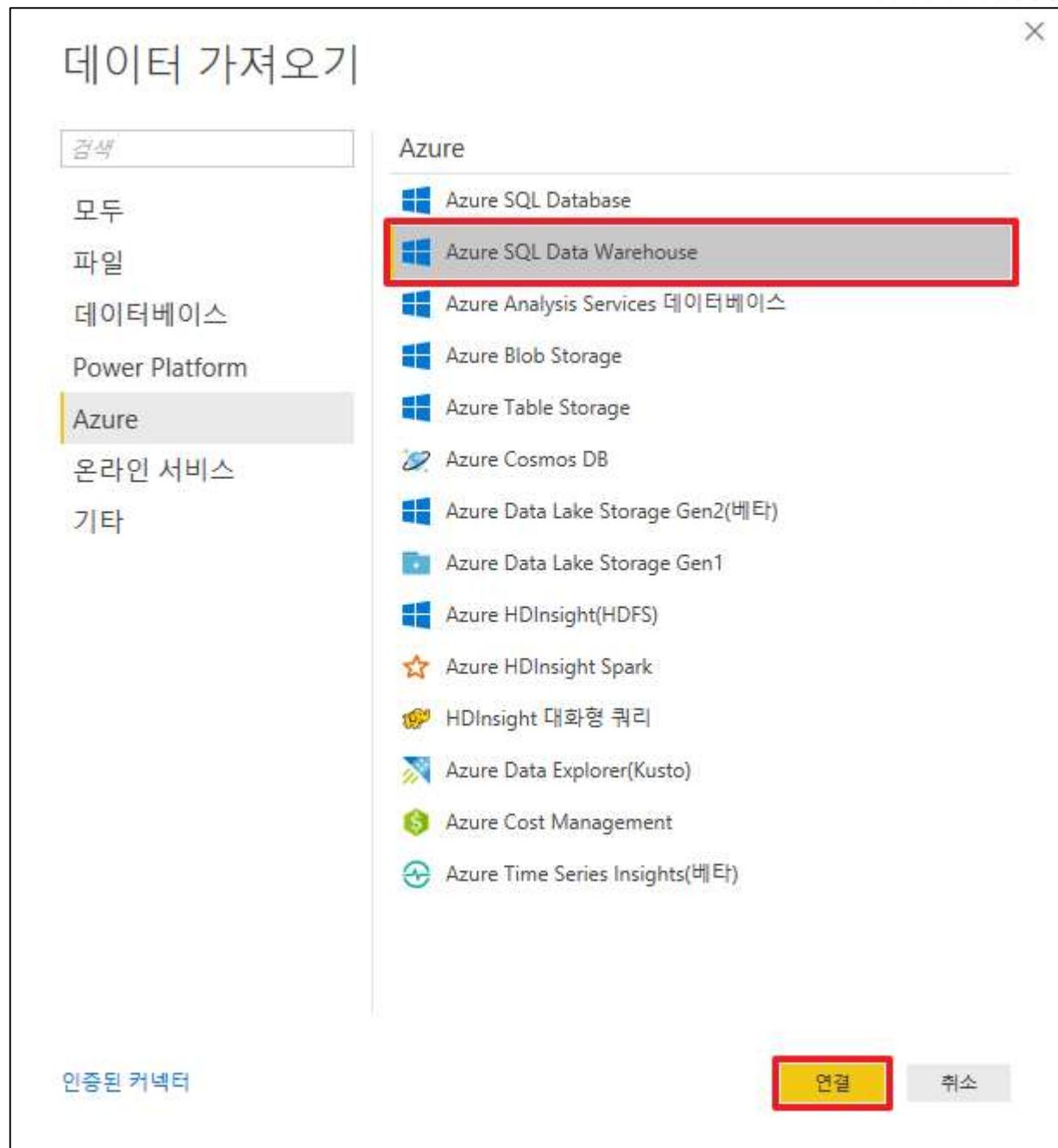


5.2. Task 2: Query data with Power BI Desktop

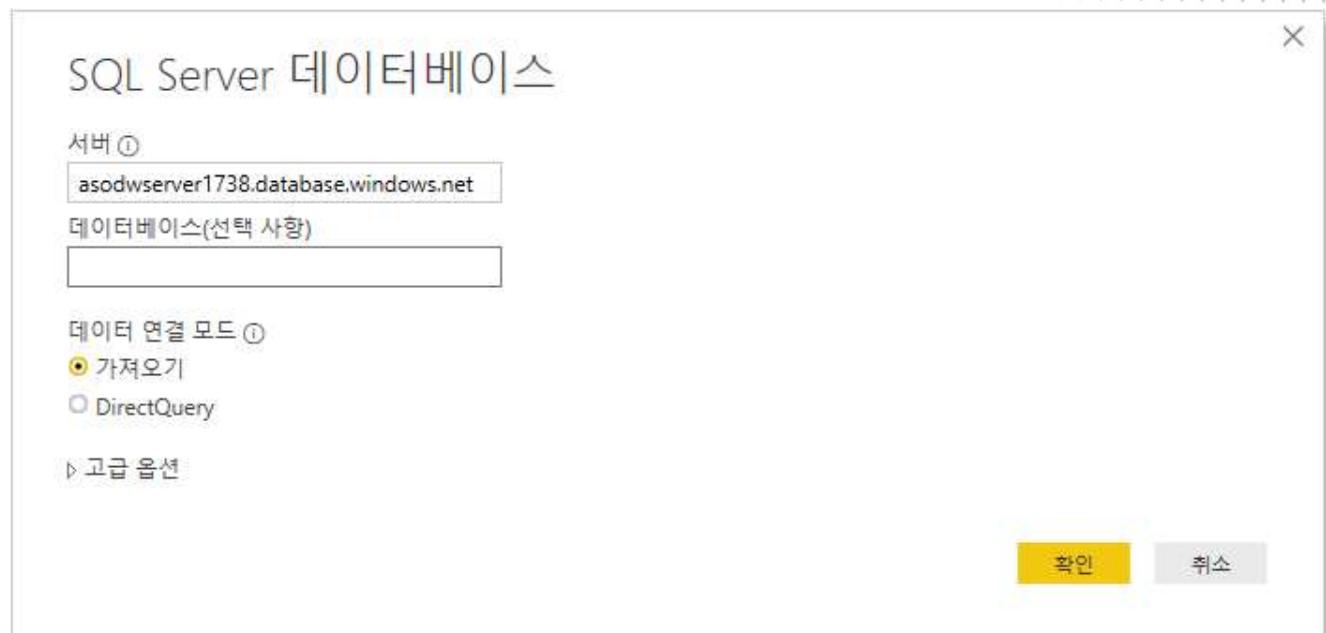
- Power BI에서 데이터 가져오기 버튼을 클릭한 후 추가...을 선택합니다.



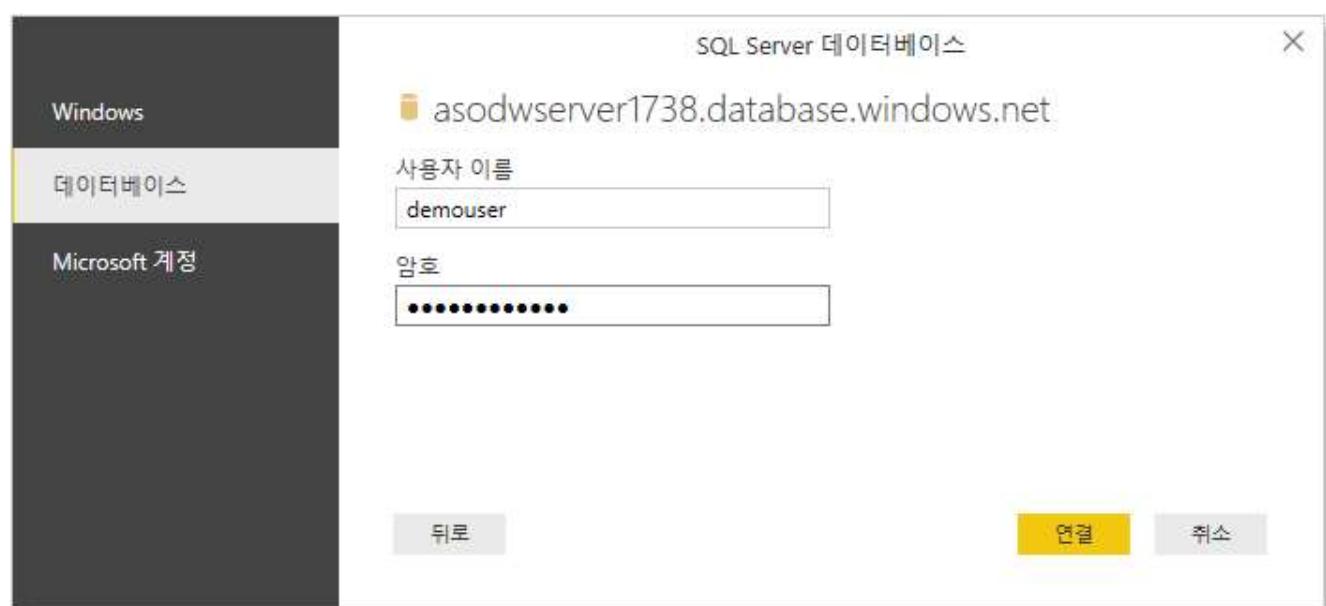
- 데이터 가져오기 창에서 왼쪽 목록에서 Azure를 선택합니다. 그런 다음 Azure SQL Data Warehouse를 선택하고 연결을 클릭합니다.



3. SQL Server 데이터베이스 화면에서 Azure Data Warehouse 서버의 이름을 입력, 데이터 연결 모드는 가져오기를 선택한 다음 확인을 클릭합니다.



4. 데이터베이스를 선택한 후 Data Warehouse의 사용자 이름과 암호를 입력한 후 **연결**을 클릭합니다.



5. 모든 테이블을 선택한 후 **로드** 버튼을 클릭합니다.

탐색 창

The screenshot shows the Power BI Data Explorer interface. On the left, there's a tree view of data sources and tables:

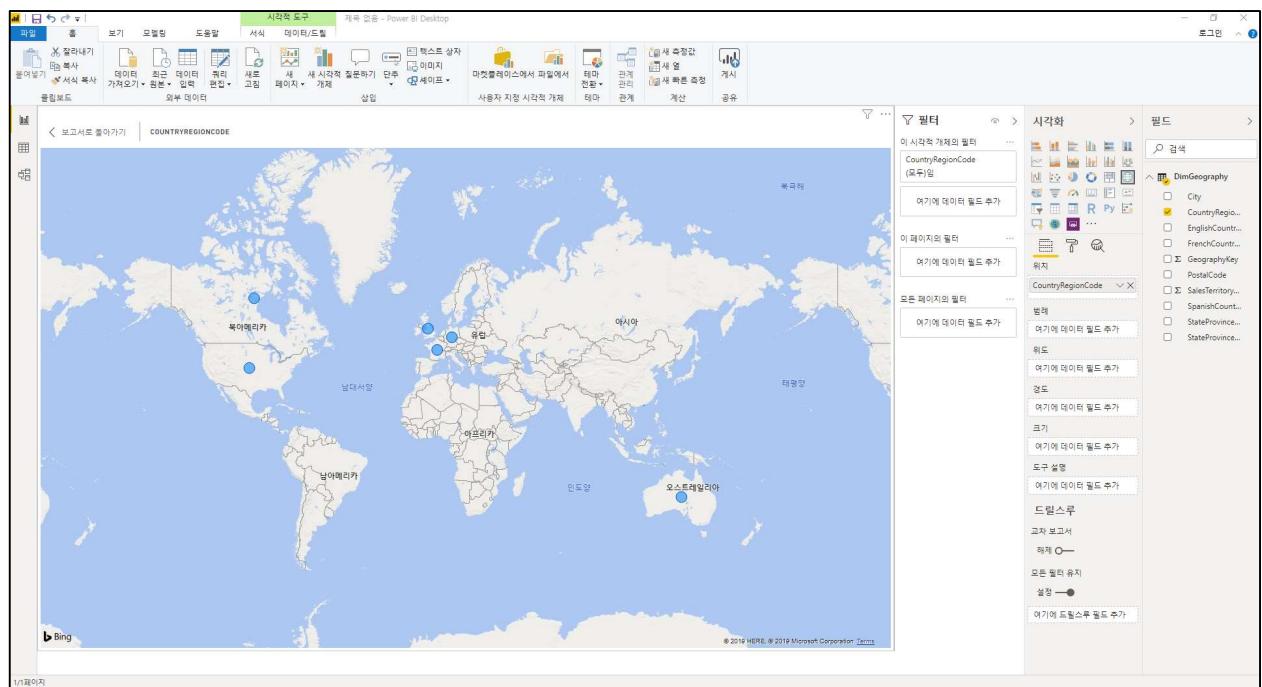
- asodwserver1738.database.windows.net [1] (selected)
- CohoDW1738 [3]
 - DimCustomer
 - DimGeography
 - FactInternetSales (selected)

On the right, the table "FactInternetSales" is displayed with the following data:

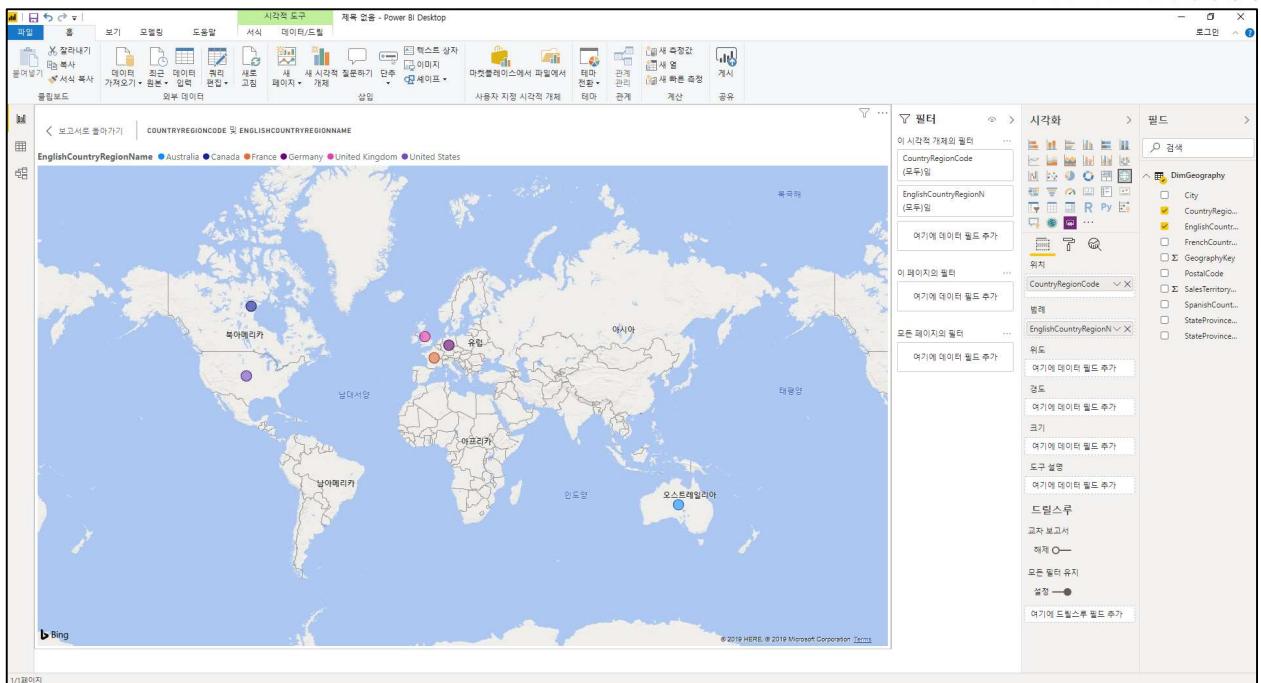
ProductKey	OrderDateKey	DueDateKey	ShipDateKey	CustomerKey	Pro
337	20021220	20030101	20021227	26547	
368	20030116	20030128	20030123	15095	
352	20030212	20030224	20030219	27050	
381	20030308	20030320	20030315	25051	
360	20030404	20030416	20030411	12854	
321	20030427	20030509	20030504	26918	
371	20030521	20030602	20030528	18160	
602	40080428	40080452	40080442	41356	
957	40080436	40080602	40080450	33800	
1070	40062056	40062218	40062208	22776	
691	40062204	40062228	40062218	45896	
1340	60093315	60093351	60093336	81387	
310	20010701	20010713	20010708	21768	
310	20010810	20010821	20010826	22061	

At the bottom, there are buttons: "관련 테이블 선택" (Select Related Tables), "로드" (Load) in yellow, "데이터 변환" (Data Transformation), and "취소" (Cancel).

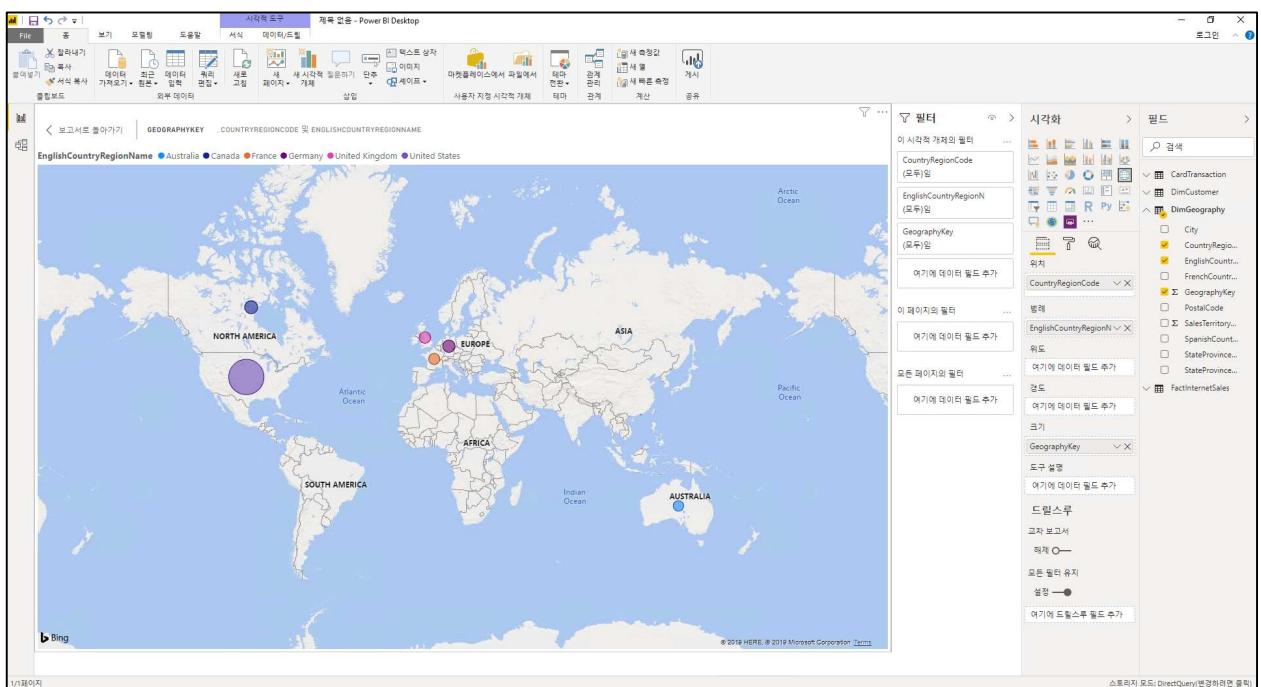
6. 오른쪽에 있는 필드 블레이드에서 DimGeography 테이블을 확장하고 CountryRegionCode 옆에 있는 상자를 선택합니다. Power BI는 지리 데이터를 인식하고 Map 시각화가 자동으로 시작됩니다.



7. 이 원이 실제로 무엇을 의미하는지에 대해 좀 더 구체적인 세부 사항을 보고 싶습니다. 국가/지역을 식별하는 범례를 추가해 봅시다. Legend 아래에서 EnglishCountryName 필드를 드래그합니다.



8. 원의 크기를 변경시켜 보겠습니다. FactInternetSales 테이블을 확장하고 수치를 가진 필드인 SalesAmount를 드래그해서 Map에 추가해봅니다.



실습이 완료되었습니다. 이번 실습을 통해 온-프레미스 Data Warehouse를 Azure SQL Data Warehouse로 마이그레이션 하는 과정에 대해 알아보았습니다

- ✓ Data Warehouse 마이그레이션 계획, 사내 구축 환경의 데이터베이스에서 클라우드로의 효율적인 마이그레이션
- ✓ Azure SQL Data Warehouse의 데이터 배포, ETL 작업을 Azure Data Factory에 연결

금일 교육에 사용된 파일들은 <https://github.com/azure-datasolution/ASO> 를 통해 다운로드 받으실 수 있습니다.