



Azure Service

ASO Hands on Lab

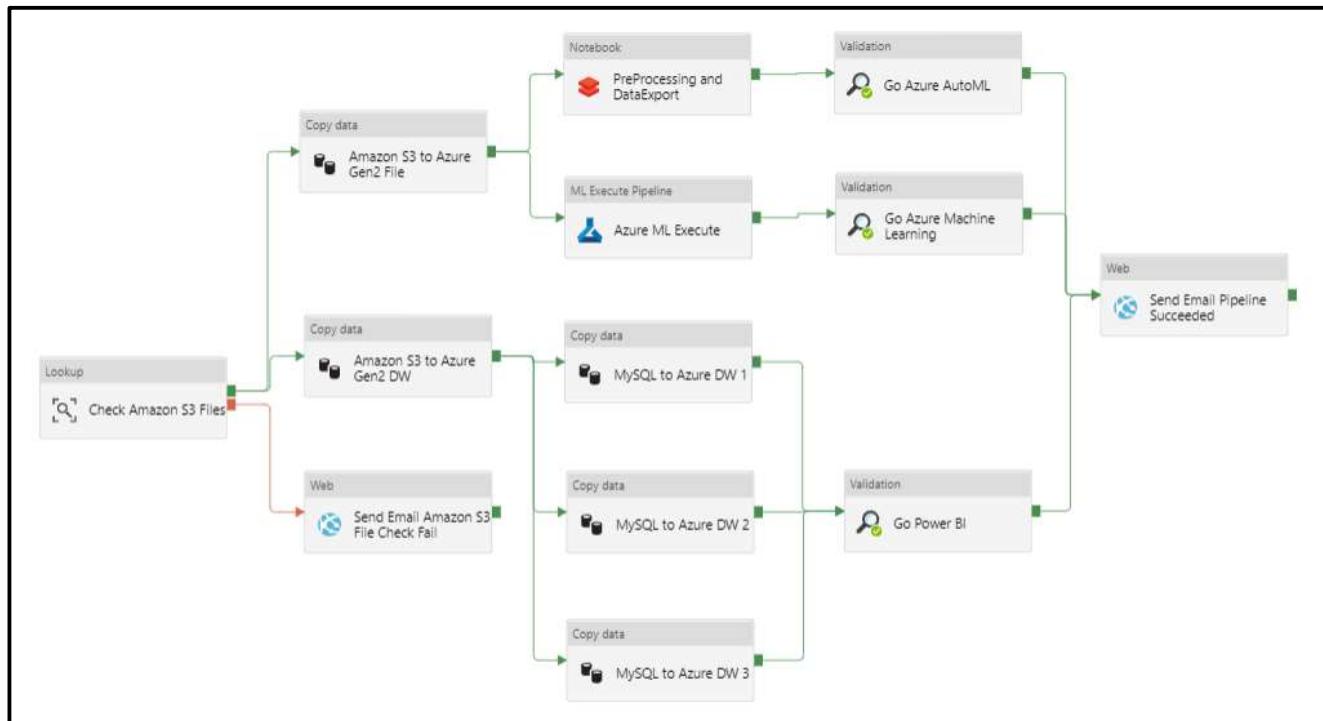
Contents

1. 사전 준비	2
1.1. 시나리오 개요	2
1.2. Azure 서비스 및 관련 제품	2
1.3. 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	10
3.4. Task 4: Create an Azure Data Factory v2	14
4. Exercise 2: Orchestrate with Azure Data Factory.....	16
4.1. Task 1: Amazon S3 to Azure Gen2.....	16
4.2. Task 2: MySQL to Azure Synapse Analytics.....	26
4.3. Task 3: Azure ML Execute	44
4.4. Task 4: PreProcessing and DataExport	46
4.5. Task 5: Send Email	48
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

이 실습은 Azure Data factory를 이용하여 다양한 데이터를 Data Lake Storage Gen2에 수집, DW Data는 Azure Data Warehouse에 적재하여 Power BI를 통한 시각화를 진행하고, 정형 데이터와 비정형 데이터는 Databricks를 통한 데이터 전 처리를 통해 머신 러닝에 적합한 상태로 변경하여 변경한 데이터로 Azure Machine Learning 서비스에 연결하여 처리합니다.



1.2. Azure 서비스 및 관련 제품

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

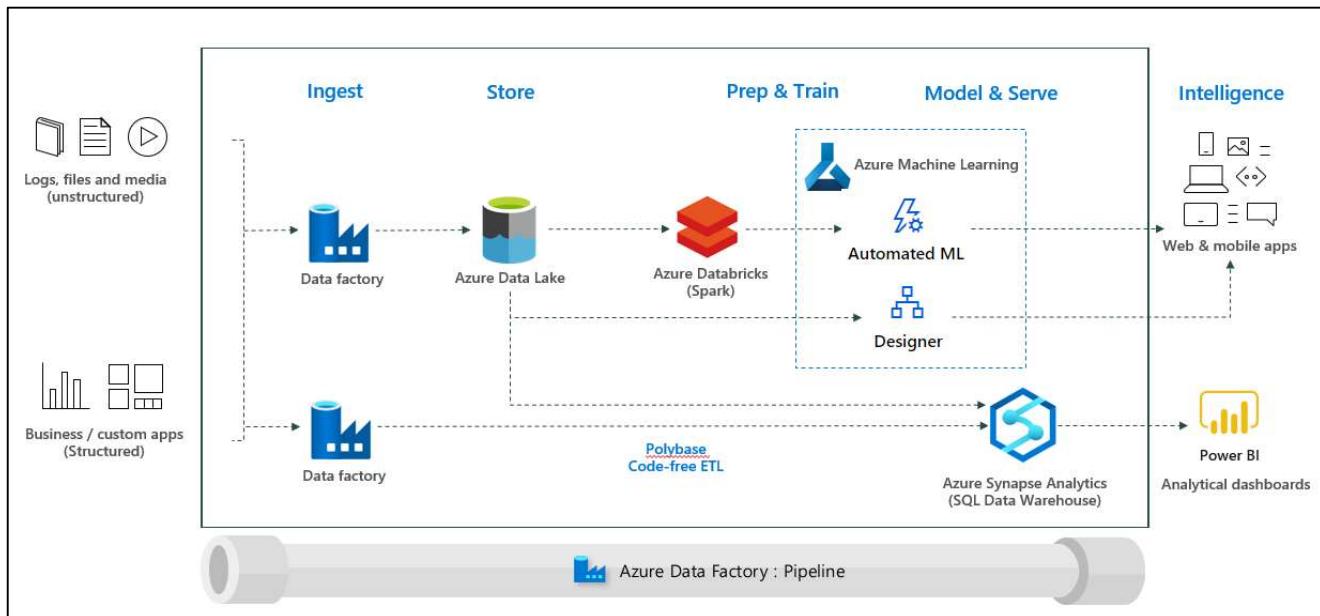
1.3. 필수 조건

- Azure Subscription

2. 학습 목표

이 실습을 통해 Azure Data factory를 이용하여 File 및 DB를 이관하는 방법과 다른 Azure 서비스와의 연계를 통해 Data를 처리하는 방법을 알아보아 Azure Data factory로 전체 Data의 흐름을 관리하는 방법을 연습해 봅니다.

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 File을 Azure 상으로 이관할 수 있도록 합니다.

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 red box highlights the 'Storage 계정 - Blob, File, Table, Queue' option, which is listed under the 'Microsoft' provider. The description indicates it's an '안정적이고 경제적인 클라우드 저장소에 Blob, 테이블, 큐 및 파일을 사용할 수 있습니다.' (A stable and cost-effective cloud storage solution for Blob, Table, Queue, and file usage).

3. 저장소 계정 블레이드에서 **만들기**를 클릭합니다. 다음 정보를 입력합니다.

- ✓ 구독: 본인의 구독 선택
- ✓ 리소스 그룹: 새로 만들기 클릭하여 생성 -> **ASO_EduXXXX**
- ✓ 스토리지 계정 이름: **asostorageXXXX**
- ✓ 위치: (아시아 태평양)한국 중부
- ✓ 성능: 표준
- ✓ 계정 종류: **StorageV2 (범용 v2)**
- ✓ Replication: **LRS(로컬 중복 스토리지)**
- ✓ 액세스 계층: **핫**

홈 > 새로 만들기 > Marketplace > Storage 계정 - Blob, File, Table, Queue > 스토리지 계정 만들기

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. 폴더 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

계정 종류 StorageV2(범용 v2)

복제 LRS(로컬 중복 스토리지)

액세스 계층(기본값) 핫

검토 + 만들기 < 이전 다음: 네트워킹 >

- 상단의 고급을 클릭하고 Data Lake Storage Gen2를 사용으로 변경합니다.

홈 > 새로 만들기 > Marketplace > Storage 계정 - Blob, File, Table, Queue > 스토리지 계정 만들기

스토리지 계정 만들기

기본 사항 네트워킹 고급 태그 검토 + 만들기

보안

보안 전송 필요 ① 사용 안 함 사용

Azure Files

대용량 파일 공유 ① 사용 안 함 사용

데이터 보호

Blob 일시 삭제 ① 사용 안 함 사용

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

Data Lake Storage Gen2

계층 구조 네임스페이스 ① 사용 안 함 사용

NFS v3 ① 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.

The screenshot shows the Azure Storage Account settings page for 'asostorage0000'. On the left, there's a sidebar with options like '개요', '활동 로그', '액세스 제어(IAM)', etc. The main area displays account details: 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. Below these details is a section titled '컨테이너' (Container) with a sub-section '대규모로 확장 가능한 데이터 레이크 스토리지' (Large-scale data lake storage). A red box highlights this '컨테이너' section.

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

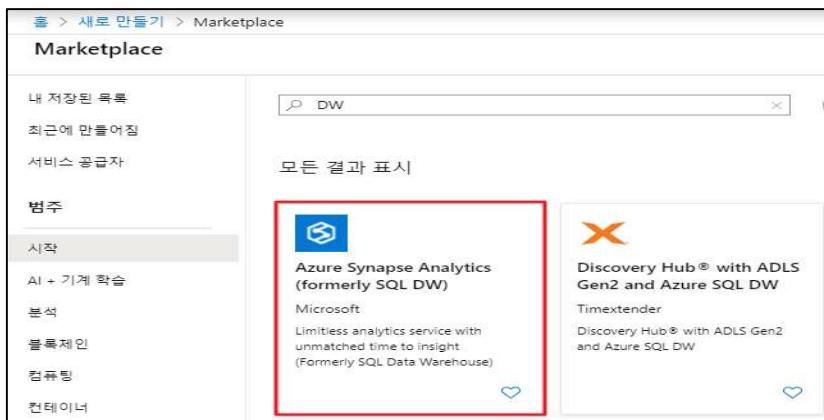
The screenshot shows the 'Containers' creation dialog for 'asostorage0000'. It has fields for '이름*' (Name*) containing 'demo' and '공용 액세스 수준' (Access level) set to '프라이빗(익명 액세스 없음)' (Private (Anonymous access off)). At the bottom are '확인' (Confirm) and '취소' (Cancel) buttons. A red box highlights the '+ Container' button in the top navigation bar.

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

The screenshot shows the 'Containers' list page for 'asostorage0000'. It lists two containers: 'demo' and 'structured'. The 'demo' container is selected, as indicated by a checkmark in its checkbox column. The 'structured' container is listed below it with its creation date and access level. The top navigation bar shows the '+ Container' button.

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

Azure Synapse Analytics
Microsoft

기본* **추가 설정*** **태그** **검토 + 만들기**

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

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. 풀더 같은 리소스 그룹을 사용하여 모든 리소스를 정리 및 관리합니다.

구독 * Azure subscription 1

리소스 그룹 * ASO_Edu0000

[새로 만들기](#)

데이터 웨어하우스 세부 정보

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

데이터 웨어하우스 이름 * CohoDW0000

서버 * (새로 만들기) asodwserver0000 ((아시아 태평양)한국 중부)

[새로 만들기](#)

성능 수준 * Gen2
DW1000c
[성능 수준 선택](#)

검토 + 만들기 **다음: 추가 설정 >**

4. 서버에서 새로 만들기를 클릭하여 아래 내용을 입력 후 확인 버튼을 클릭합니다.

- ✓ 데이터 웨어하우스 이름: CohoDWXXXX
- ✓ 서버 이름: asodwserverXXXX (DW 계정 명)
- ✓ 서버 관리자 로그인: demouser
- ✓ 암호: Demo@pass123
- ✓ 위치: (아시아 태평양)한국 중부
- ✓ Azure 서비스의 서버 액세스 허용: 체크함



5. 성능 수준 선택을 클릭하고 Gen2 탭을 선택한 다음 성능을 DW100c로 설정하고 적용을 클릭합니다.



6. 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..

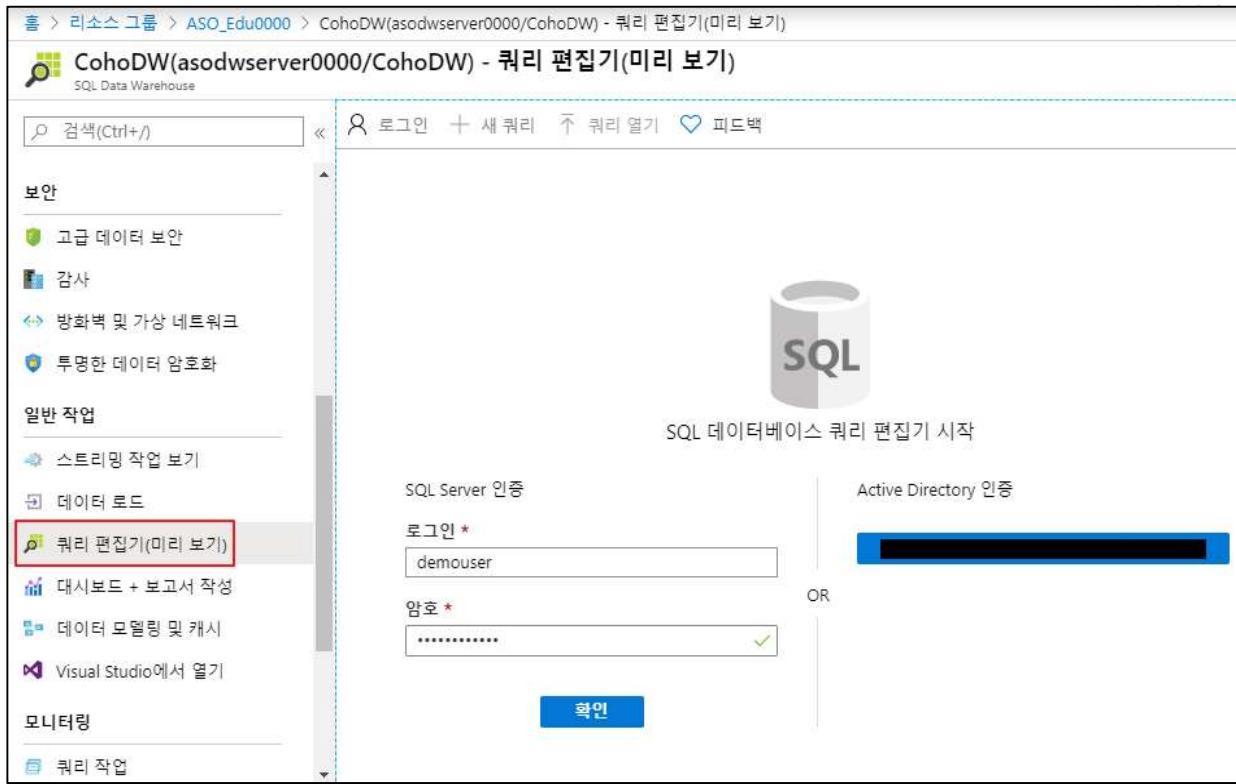
The screenshot shows the Azure Resource Group 'ASO_Edu0000'. On the left, there's a sidebar with icons for Overview, Activity Log, Access Control (IAM), Tags, and Events. Below that is a 'Setting' section with icons for Quick Start, Resource Usage, Metrics, Policies, and Metrics. The main area lists resources under '구독 (변경) : Azure subscription 1'. It includes 'Azure Data Factory V2' (Data Factory(V2)), 'SQL Server' (SQL Server), 'Storage Account' (스토리지 계정), and 'SQL Data Warehouse' (SQL Data Warehouse). The 'CohoDW(asodwserver0000/CohoDW)' resource is highlighted with a red box at the bottom of the list.

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. On the left, there's a sidebar with icons for Connection Strings, Metrics, Firewall & Virtual Network, Advanced Data Protection, and Auditing. The main area has a '저장' (Save) button highlighted with a red box. Below it is a '클라이언트 IP 추가' (Add Client IP) button also highlighted with a red box. A note says '아래 지정된 IP 연결에서 asodwserver0000의 모든 데이터베이스에 대한 액세스 권한을 제공합니다.' (IP address specified below provides access to all databases on asodwserver0000). It shows a table with a single entry: '클라이언트 IP 주소' (Client IP Address) is '223.62.175.102', '규칙 이름' (Rule Name) is 'ClientIPAddress_2020-2-20...', '시작 IP' (Start IP) is '223.62.203.89', and '종료 IP' (End IP) is '...'. The '방화벽 및 가상 네트워크' (Firewall & Virtual Network) section is also highlighted with a red box at the bottom.

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

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



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)

```

The screenshot shows the CohoDW application interface. On the left, there's a sidebar with a user icon, navigation links like '테이블' (Table), '뷰' (View), and '저장 프로시저' (Save Procedure), and a message box stating '여기에서 제한된 개체 탐색기를 표시하는 중입니다. 전체 기능을 확인하려면 SSDT를 여세요.' (A limited object browser is displayed here. To view all features, open SSDT). The main area has a title '쿼리 1' (Query 1) with a red box around the '실행' (Execute) button. Below it is a code editor with a SQL script to create a 'DimCustomer' table. At the bottom, tabs for '결과' (Results) and '메시지' (Messages) are shown, with the message tab currently active.

```
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(20) null,`15 [EducationLevel] nvarchar(20) null,`16 [AnnualPurchases] smallint null,`17 [AnnualSales] money null,`18 [AvgOrderPrice] money null,`19 [CustomerDemographicKey] int null,`20 [CustomerSegmentKey] int null,`21 [LastEditedBy] nvarchar(16) null,[ModifiedDate] datetime not null)
```

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

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

결과	메시지			
<input type="text"/> 필터링 항목 검색...				
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.

The screenshot shows the Azure Marketplace search results for 'factory'. On the left, there's a sidebar with categories like '내 저장된 목록', '최근에 만들어짐', '서비스 공급자', '범주', and '시작'. The '시작' category is selected. In the main area, there's a search bar with 'factory' typed in. Below it, there are two items: 'Data Factory' by Microsoft and 'VIACode Azure Data Factory Monitor' by VIACode. The 'Data Factory' item is highlighted with a red box.

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
- ✓ Git 사용 미 체크

The screenshot shows the '새 data factory' creation form. It includes fields for '이름' (Name) set to 'ASODatafactory0000', '버전' (Version) set to 'V2', '구독' (Subscription) set to 'Azure subscription 1', '리소스 그룹' (Resource Group) set to 'ASO_Edu0000', '새로 만들기' (Create New), '위치' (Location) set to '한국 중부', and 'GIT 사용' (Use GIT) with an unchecked checkbox. At the bottom is a '만들기' (Create) button.

4. 만들기 버튼을 클릭합니다.

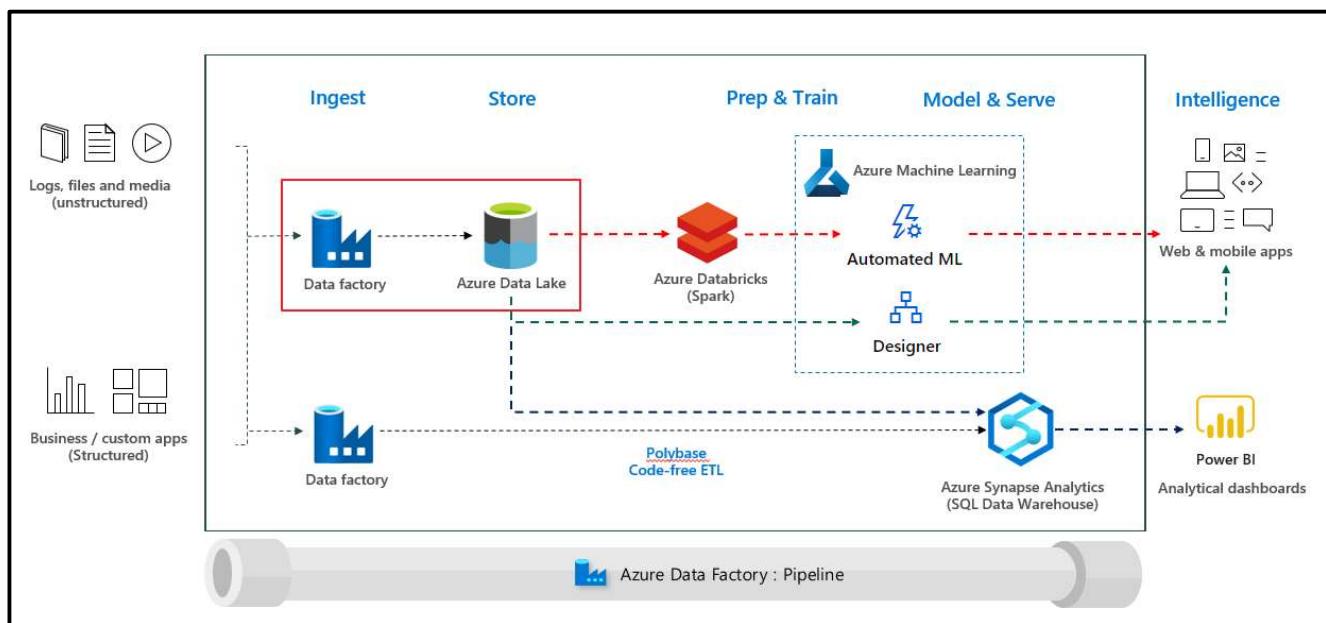
Data Factory 위치는 Data Factory의 메타데이터가 저장되고 파이프라인의 트리거가 시작되는 위치입니다. 한편, Data Factory는 다른 Azure 영역의 데이터 저장소 및 컴퓨팅 서비스에 액세스하여 데이터 저장소 간에 데이터를 이동하거나 컴퓨팅 서비스를 사용하여 데이터를 처리할 수 있습니다. 이 동작은 전 세계적으로 사용 가능한 통합 런타임을 통해 실현되어 데이터 컴플라이언스, 효율성 및 네트워크 송신 비용 절감을 보장합니다.

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

이 작업은 Amazon S3에 파일이 존재하는지 확인한 후 해당파일을 Azure 서비스에서 사용하기 위해 Azure Data Lake Store Gen2로 복사합니다.



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

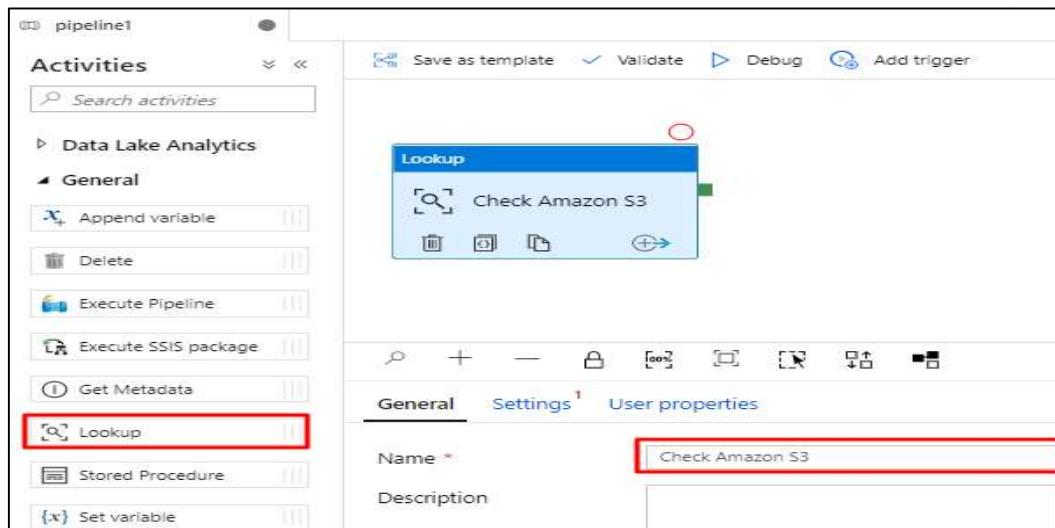
The screenshot shows the Azure portal interface for managing a Data Factory resource named 'ASODatafactory0000'. The left sidebar lists navigation options: '개요' (Overview) is selected, followed by '활동 로그' (Activity Log), '액세스 제어(IAM)', '태그', '문제 진단 및 해결', '설정', '잠금', '일반', '환경 속성', and '시작'. The main content area displays resource details: '리소스 그룹 (변경)' is 'ASO_Edu0000'; '상태' is 'Succeeded'; '위치' is '한국 중부'; '구독 (변경)' is 'Azure subscription 1'; '구독 ID' is '35120fde-ae11-40e4-8ea2-8be769fdcc5d'. On the right, there are two buttons: '문서' (Document) and '작성 및 모니터링' (Create and Monitoring). The '작성 및 모니터링' button is highlighted with a red box.

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

The screenshot shows the Azure Data Factory home page with the title 'Let's get started'. It features four main buttons: 'Create pipeline' (highlighted with a red box), 'Create data flow', 'Create pipeline from template', and 'Copy data'. Each button has a corresponding icon: a blue cylinder with a yellow plus sign for 'Create pipeline', two blue cylinders with a yellow arrow for 'Create data flow', a blue cylinder with a yellow square for 'Create pipeline from template', and two blue cylinders with yellow stars for 'Copy data'.

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.

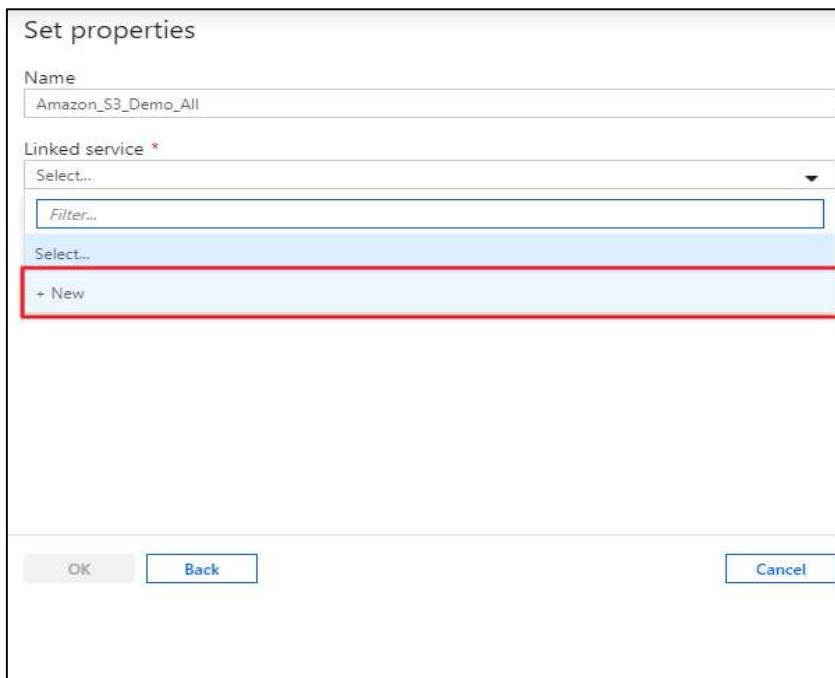


6. format은 DelimitedText를 선택합니다.



7. Name 항목 입력 후 Linked service에서 +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 : OOr4sejC3ICrmT7Nuv4RaOezH7UDatbh4sl+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
Secret Access Key *	
.....	

Service URL

Test connection To linked service To file path

Annotations

Create Connection successful Test connection **Cancel**

9. Set Properties 블레이드가 열리면 File path 항목의 Browse 버튼을 클릭하여 파일 경로를 **asodemo > Demo_All** 지정 후 OK 버튼을 클릭하여 완료합니다.
- ✓ Name : **Amazon_S3_Demo_All**
 - ✓ Linked service: **AmazonS3**
 - ✓ File path : **asodemo > Demo_All**

Set properties

Name Amazon_S3_Demo_All

Linked service * AmazonS3

Edit connection

File path asodemo / Demo_All / File **Browse** ▼

First row as header

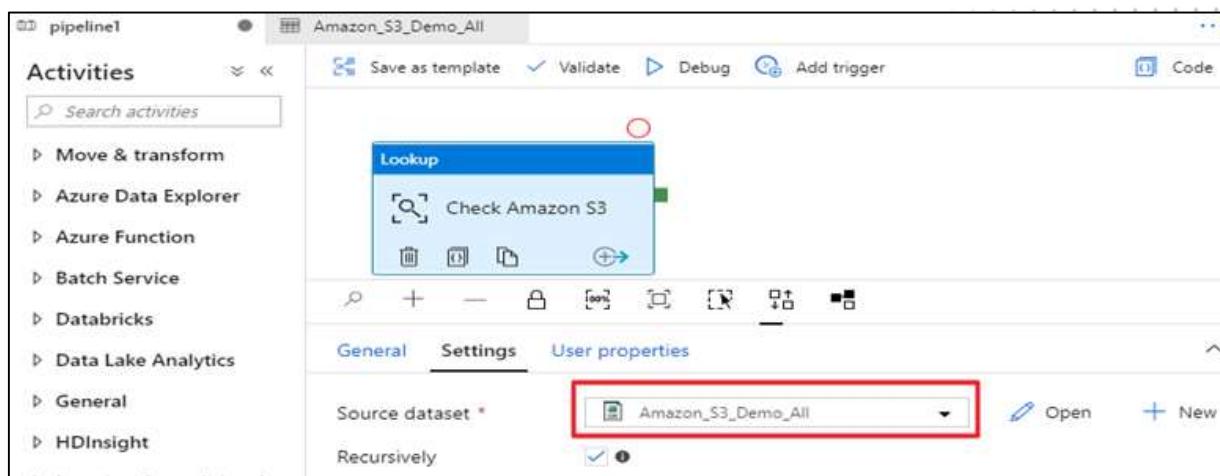
Import schema

From connection/store From sample file None

Advanced

OK **Back** **Cancel**

10. Lookup 액티비티의 **Setting**탭의 **Source Dataset**이 세팅된 것을 확인합니다.

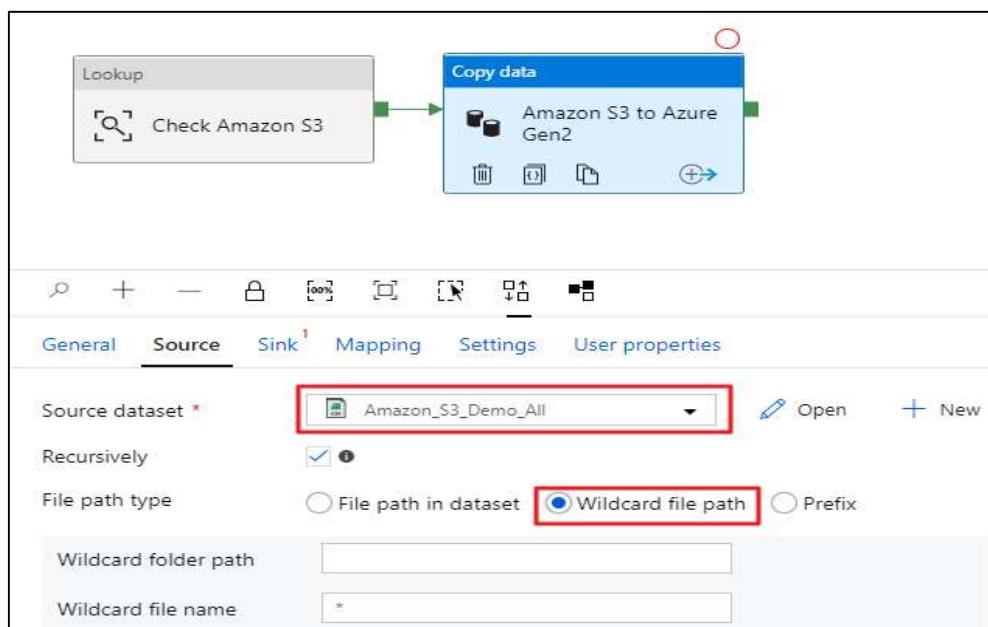


11. Activities > Move & transform > Copy data를 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.

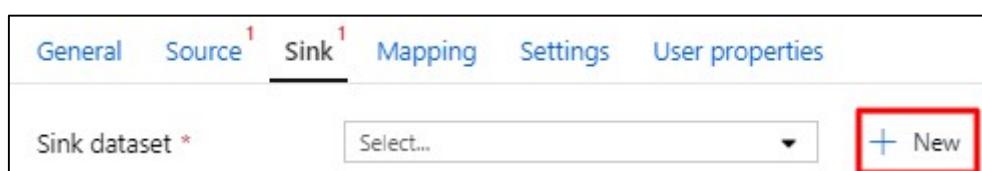
- ✓ Name: Amazon S3 to Azure Gen2

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

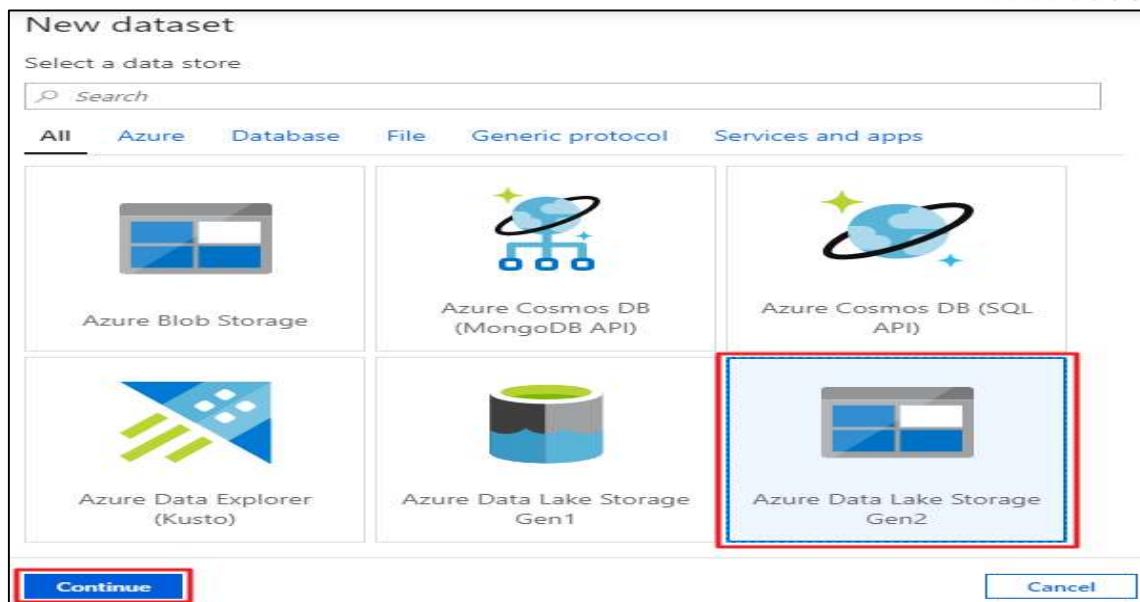
- ✓ Source dataset : Amazon_S3_Demo_All
- ✓ File path Type : Wildcard file path



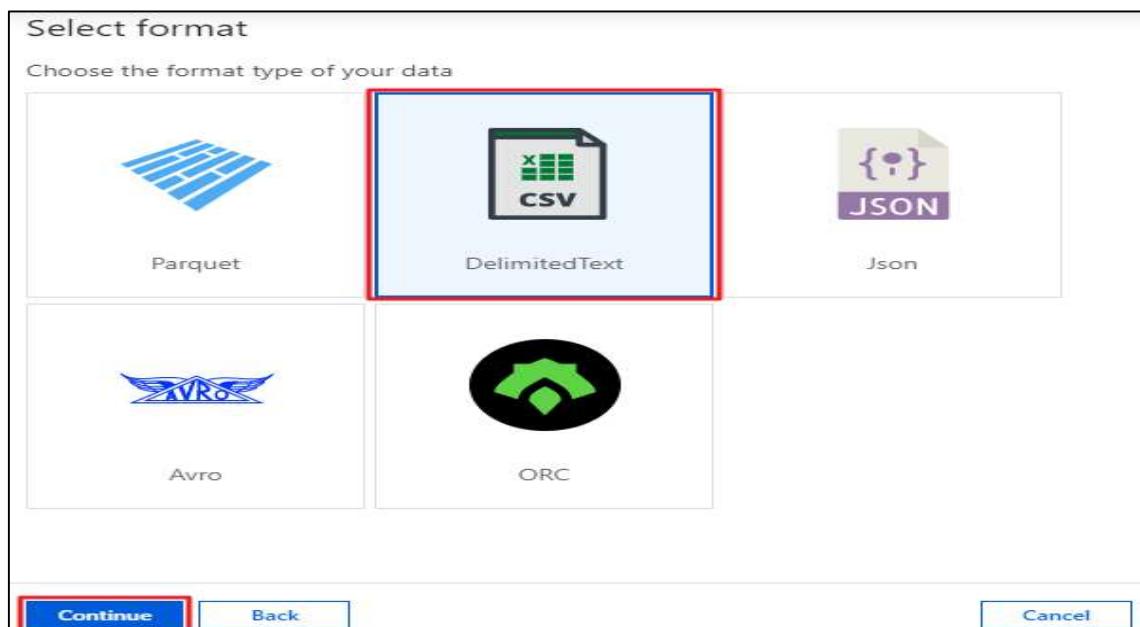
13. Sink탭으로 이동하여 Sink dataset의 +New를 클릭합니다.



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

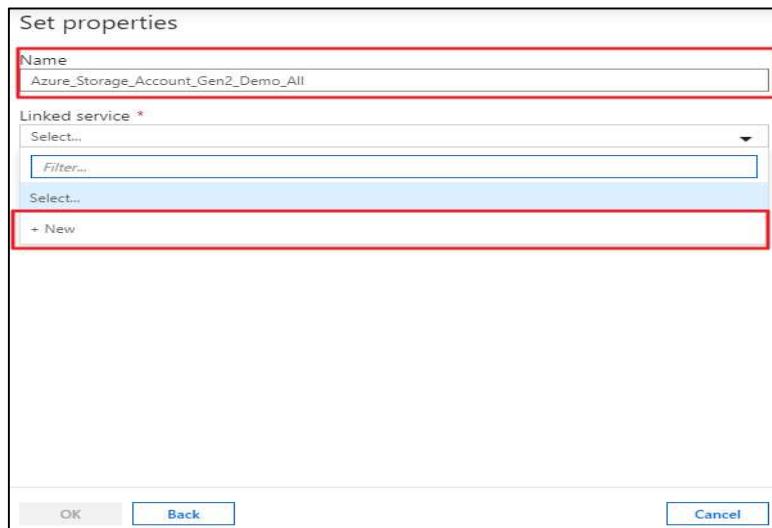


15. format은 DelimitedText를 선택합니다.



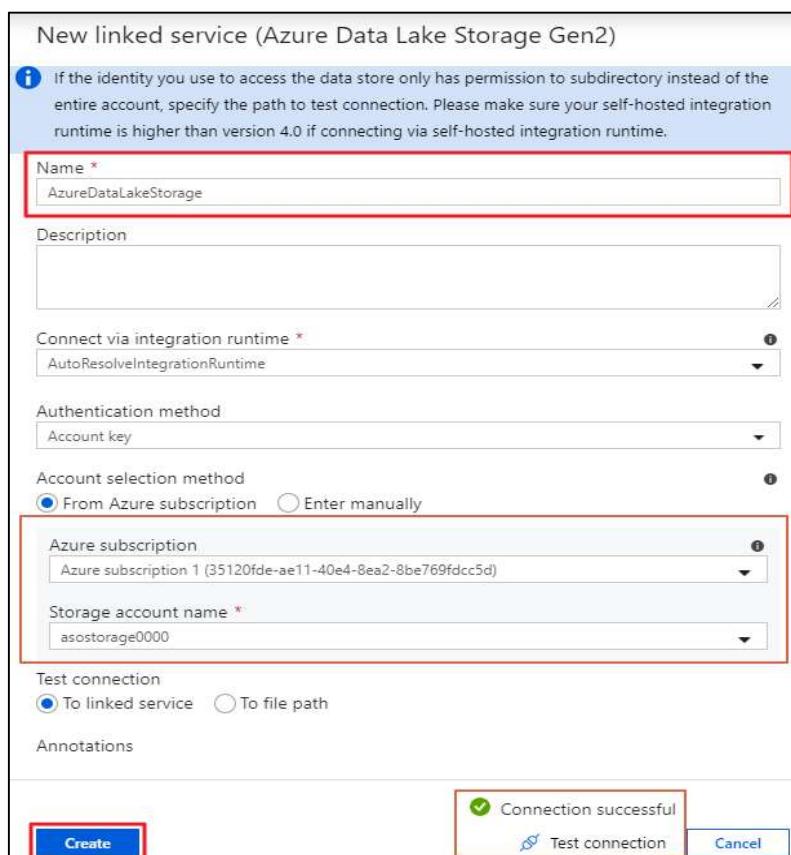
16. Set Properties 블레이드가 열리면 Name 항목을 입력 후 Linked service 의 +New 를 선택합니다.

- ✓ Name: Azure_Storage_Account_Gen2_Demo_All



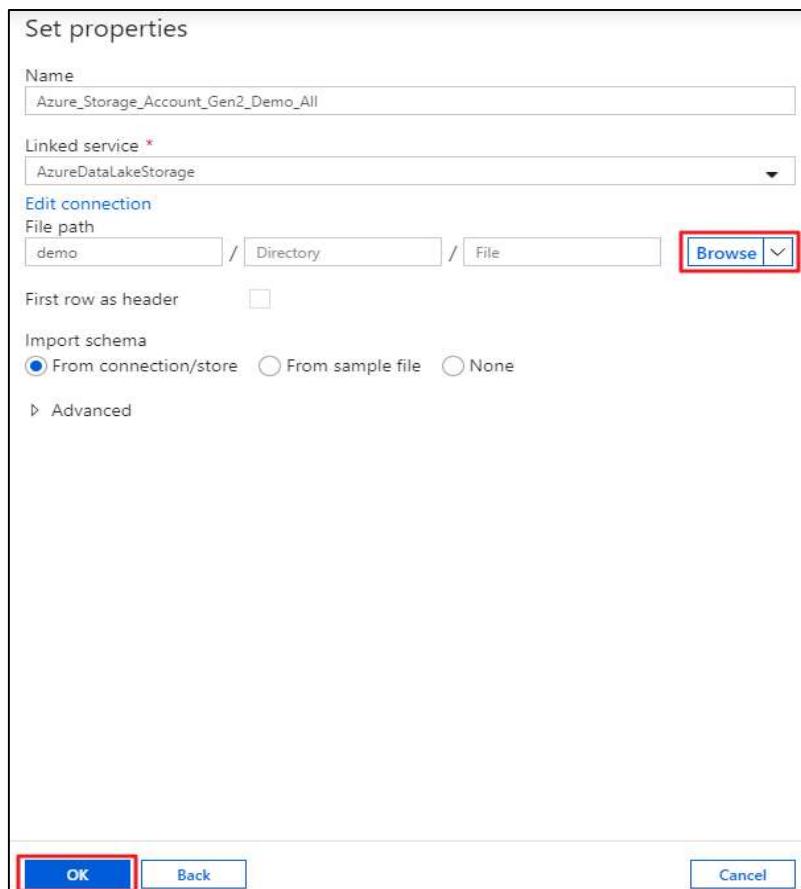
17. New linked service (Azure Data Lake Storage Gen2) 블레이드가 열리면, 아래와 같이 정보를 입력하고 Test Connection 버튼을 클릭하여 Connection successful 메시지가 출력되면 Create 버튼을 클릭합니다.

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

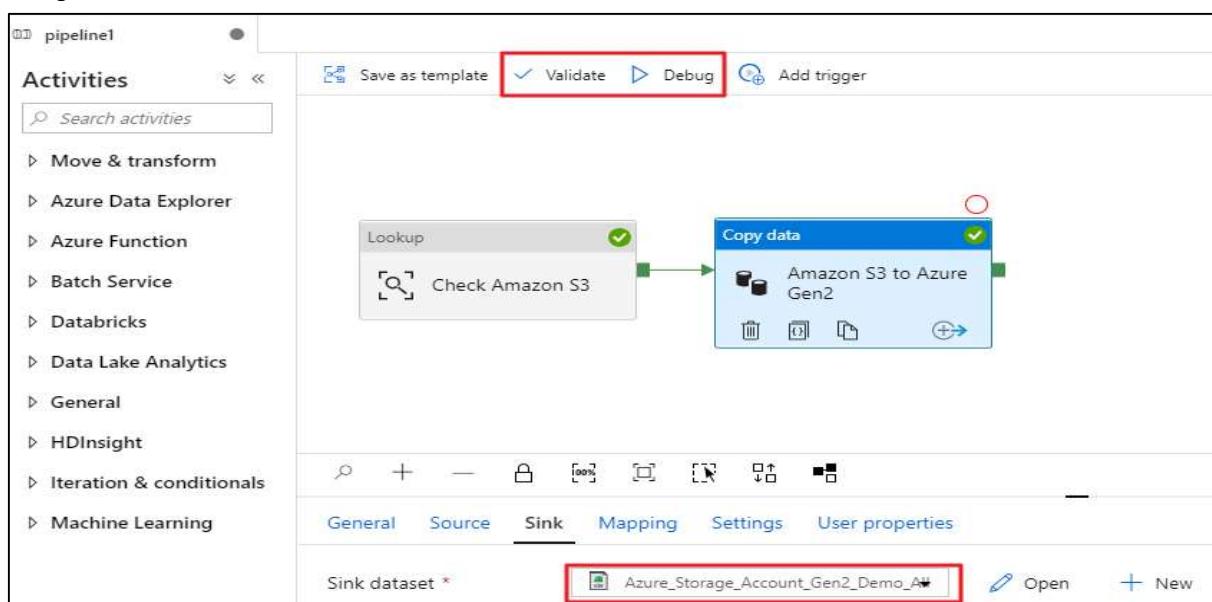


18. Set Properties 블레이드가 열리면 아래의 정보를 입력 후 OK 버튼을 클릭하여 완료합니다.

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



19. Copy data 액티비티의 Sink 탭의 Sink Dataset이 세팅된 것을 확인 후 상단의 validate 버튼을 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's Resource Groups blade for the group 'ASO_Edu0000'. On the left is a navigation menu with options like '개요', '활동 로그', '액세스 제어(IAM)', '태그', '이벤트', '설정', '리소스 비용', '배포', '정책', and '속성'. The main area displays a list of resources under the heading '구독 (변경) : Azure subscription 1'. The list includes 'ASODatafactory0000' (Data Factory(V2)), 'asodwserver0000' (SQL Server), 'asostorage0000' (Storage Account, highlighted with a red box), and 'CohoDW(asodwserver0000/CohoDW)' (SQL Data Warehouse). There are filters at the top for '이름을 기준으로 필터링...', '형식 == 모두', '위치 == 모두', and a '필터 추가' button.

21. 컨테이너 버튼을 클릭합니다.

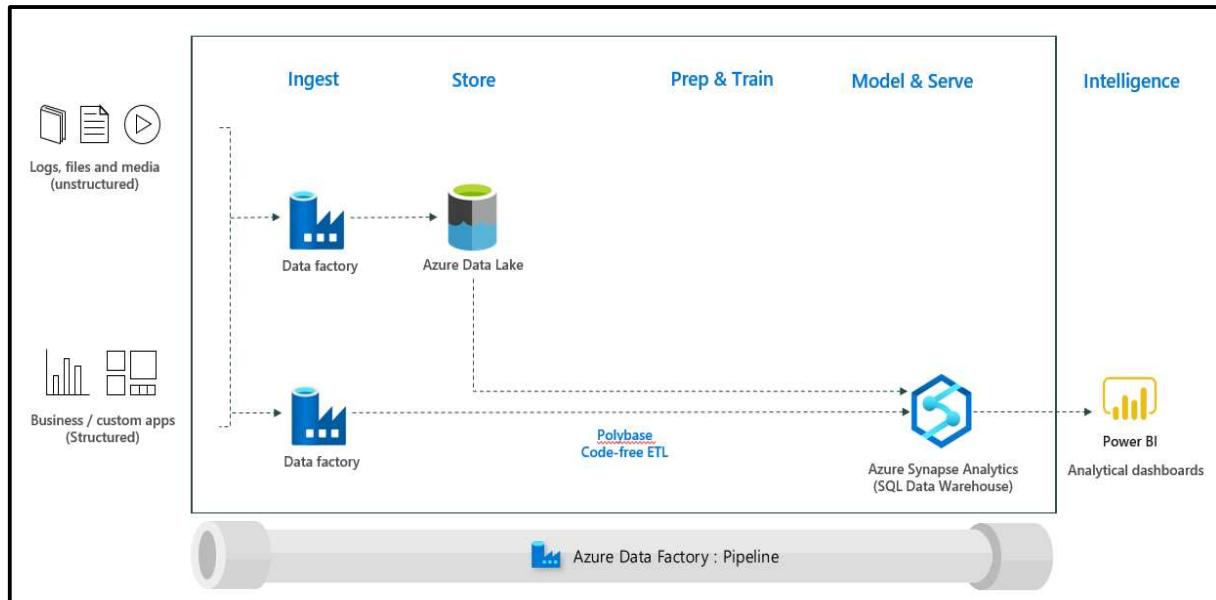
The screenshot shows the details page for the storage account 'asostorage0000'. The left sidebar has options like '개요', '활동 로그', '액세스 제어(IAM)', '태그', '문제 진단 및 해결', '데이터 전송', '이벤트', 'Storage Explorer(미리 보기)', '설정', and '액세스 키'. The main pane displays resource group information ('리소스 그룹 (변경) : ASO_Edu0000'), location ('한국 중부'), subscription ('Azure subscription 1'), and tags ('태그를 추가하려면 여기를 클릭'). Below this, there are two sections: '컨테이너' (Containers) and '파일 공유' (File sharing). The '컨테이너' section is highlighted with a red box. It contains a description of what a container is and a '자세한 정보' (More details) link.

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

The screenshot shows the details page for the storage container 'demo'. The left sidebar has options like '개요', '액세스 제어(IAM)', '설정', '액세스 정책', '속성', and '메타데이터'. The main pane shows '업로드' (Upload) and '새로 고침' (Refresh) buttons, and a note about authentication methods ('인증 방법: 액세스 키 (Azure AD 사용자 계정으로 전환)'). It also shows the location ('위치: demo') and a search bar for blob search ('접두사로 Blob 검색(대/소문자 구분)'). Below this is a table with columns '이름', '수정한 날짜', and '액세스 계층'. The table lists three items: 'mysql' (last modified 25/5/2023), 'structured' (last modified 25/5/2023), and 'unstructured' (last modified 25/5/2023).

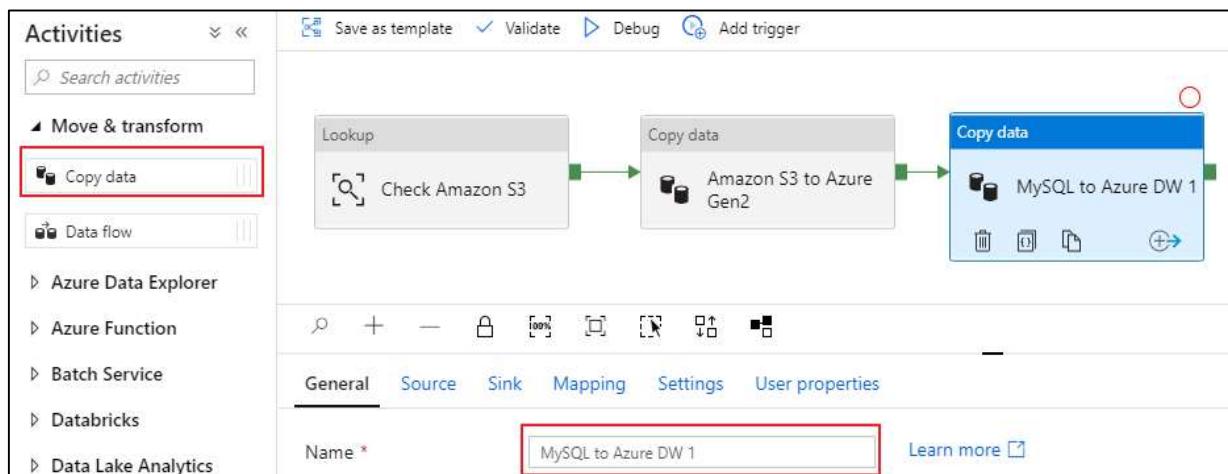
4.2. Task 2: MySQL to Azure Synapse Analytics

이 작업은 Task 1에서 복사한 파일을 이용하여 Azure Data Warehouse에 데이터를 이관합니다.

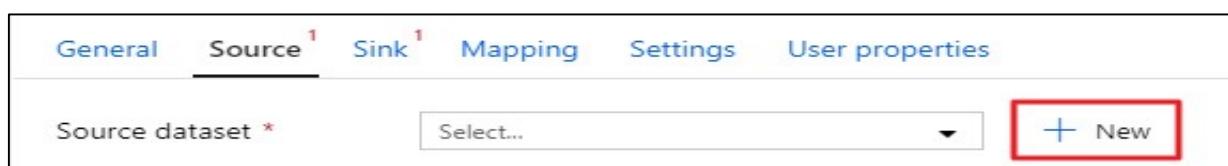


- Activities > Move & transform > Copy data를 드래그해서 옮긴 후 하단 General 탭에서 다음을 입력합니다.

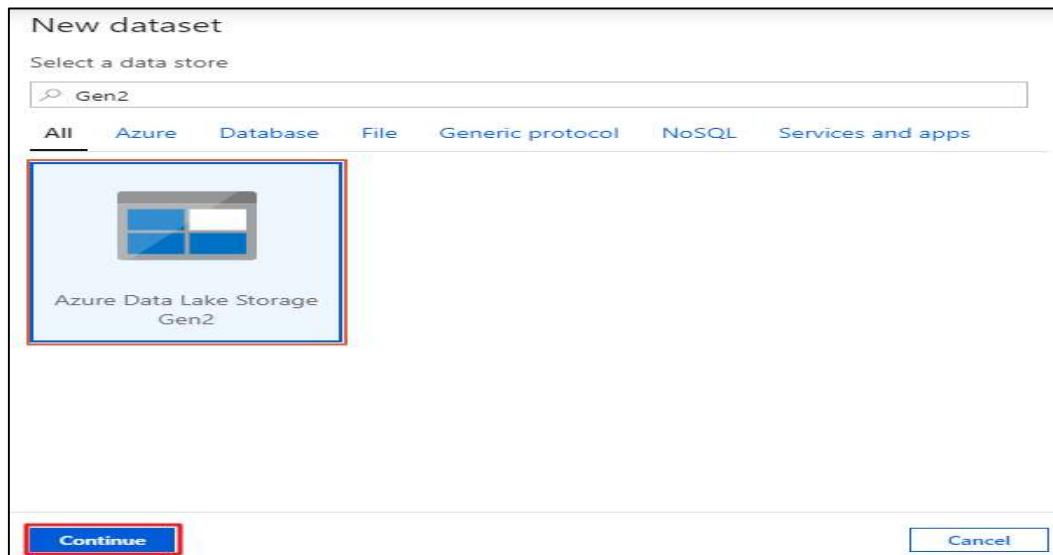
- ✓ Name: MySQL to Azure DW 1



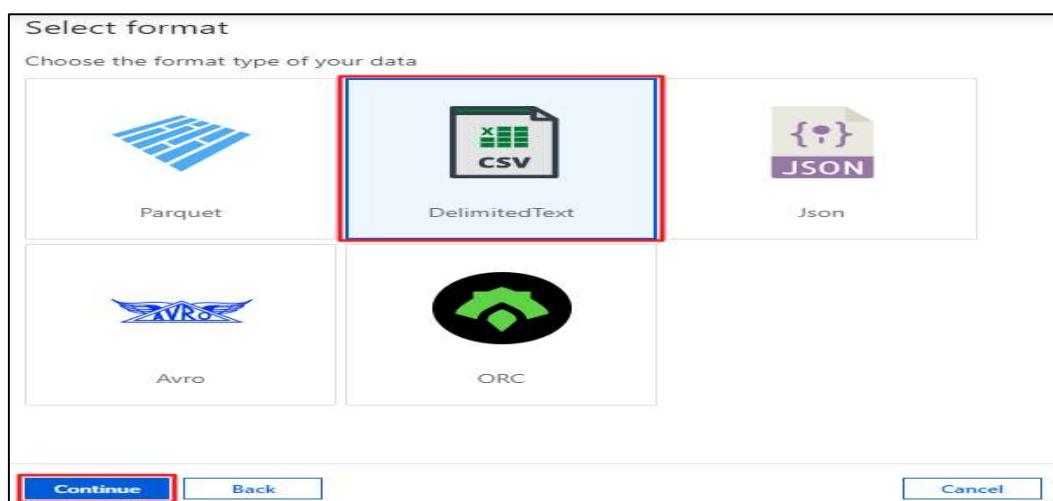
- Source 탭으로 이동하여 Source dataset의 +New 를 클릭합니다.



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

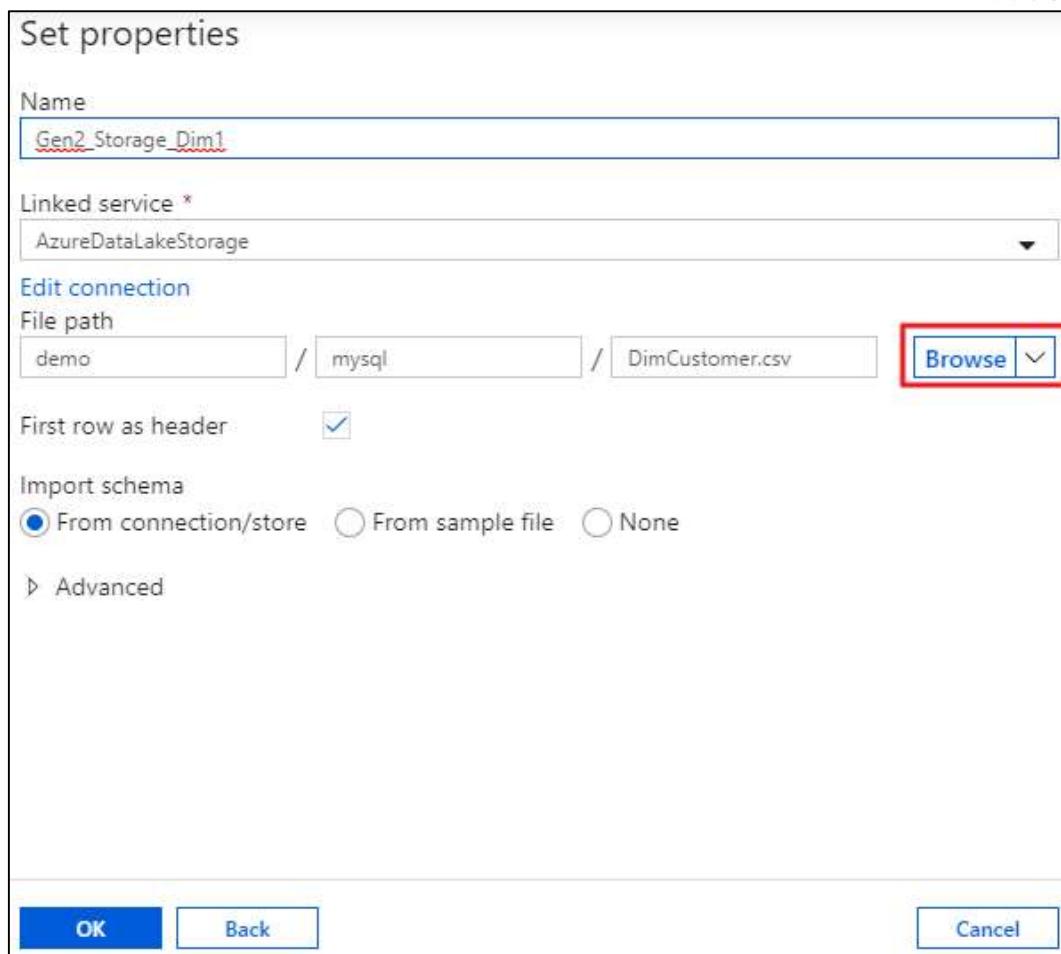


4. format은 DelimitedText를 선택합니다

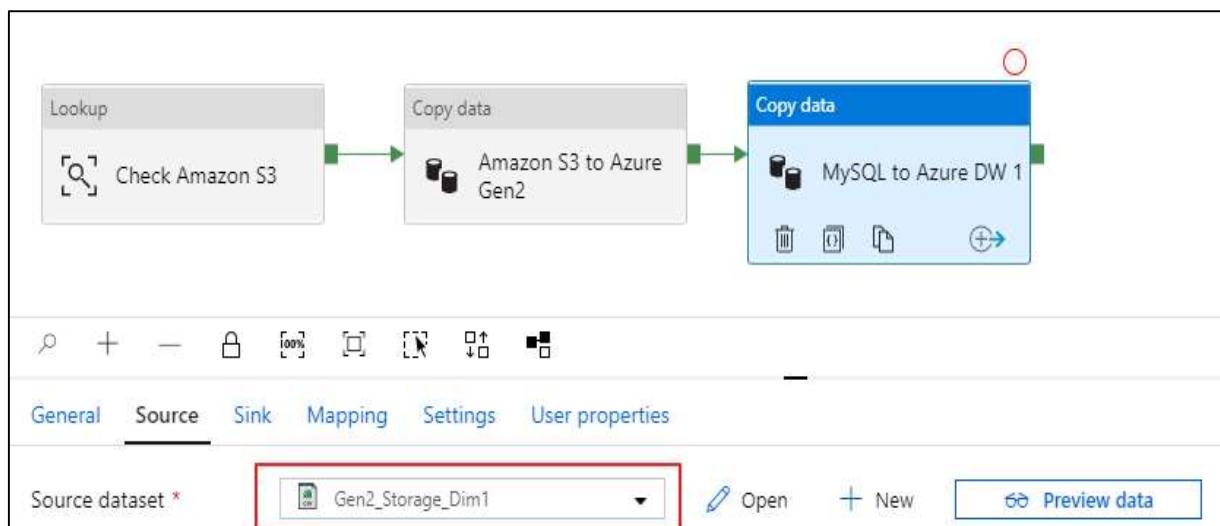


5. Set Properties 블레이드가 열리면 아래의 정보를 입력 후 OK버튼을 클릭합니다.

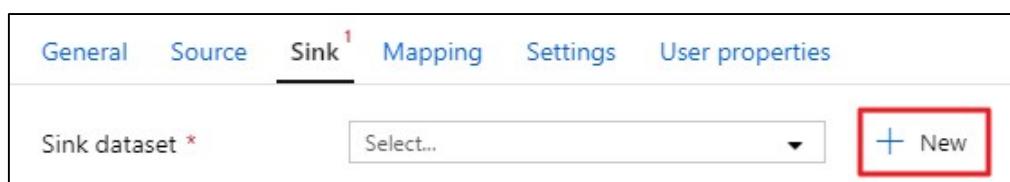
- ✓ Name: **Gen2_Storage_Dim1**
- ✓ Linked service : **AzureDataLakeStorage**
- ✓ File path: **demo / mysql / DimCustomer.csv** (browse 버튼 클릭하여 경로 지정)
- ✓ First row as header : **Check**



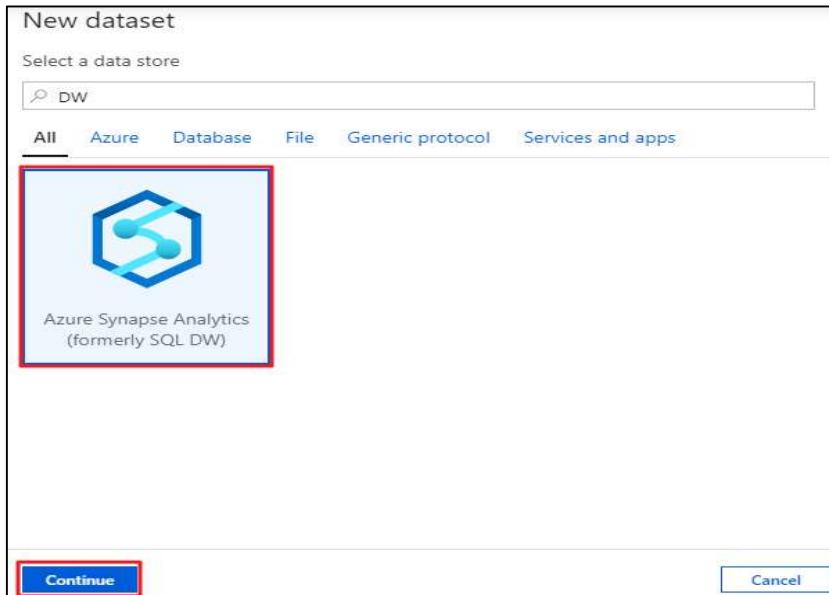
6. Source탭의 Source Dataset이 세팅된 것을 확인합니다.



7. Sink탭으로 이동하여 Sink dataset의 +New를 클릭합니다.

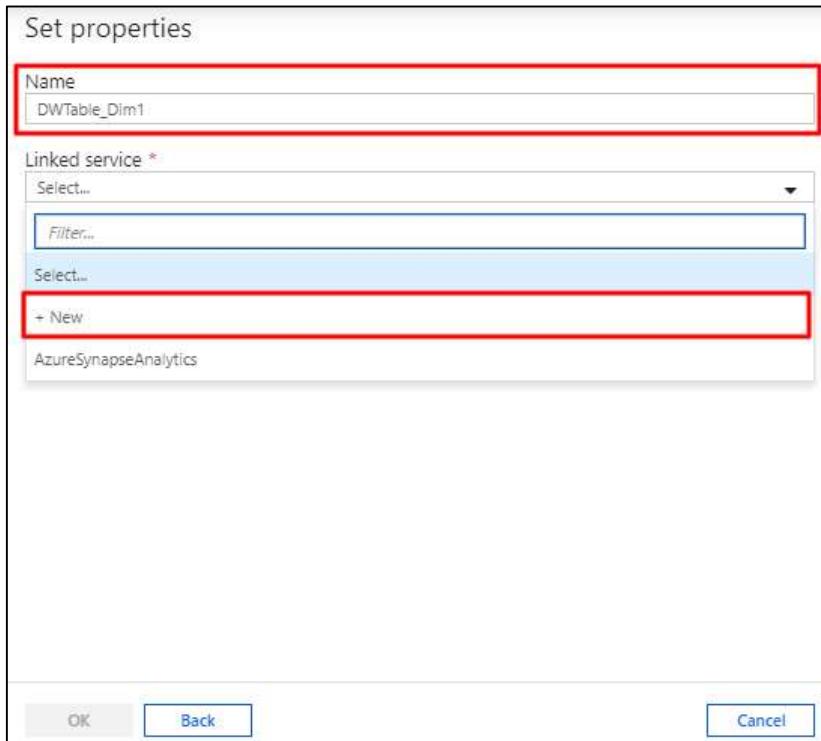


8. New dataset 블레이드에서 DW를 검색하여 Azure Synapse Analytics를 선택 후 Continue를 클릭합니다.



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

✓ Name: DWTable_Dim1



10. 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 : CohoDWXXXX
- ✓ 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 *
CohoDW0000

Authentication type *
SQL authentication

User name *
demouser

Password Azure Key Vault

Password *

✓ Connection successful 🔗 Test connection Cancel

Create Back

11. 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

12. Copy data 액티비티의 Sink탭의 Sink Dataset이 세팅된 것을 확인합니다.

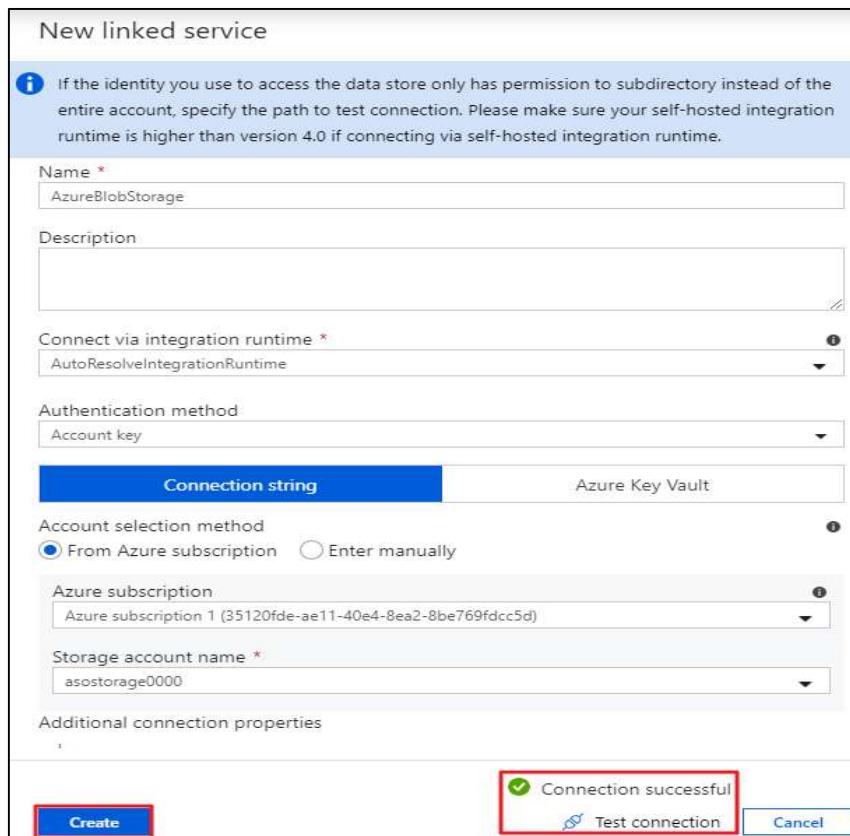
General	Source	Sink	Mapping	Settings	User properties
Sink dataset * <input type="button" value="DWTable_Dim1"/> <input type="button" value="Open"/> <input type="button" value="New"/>					

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

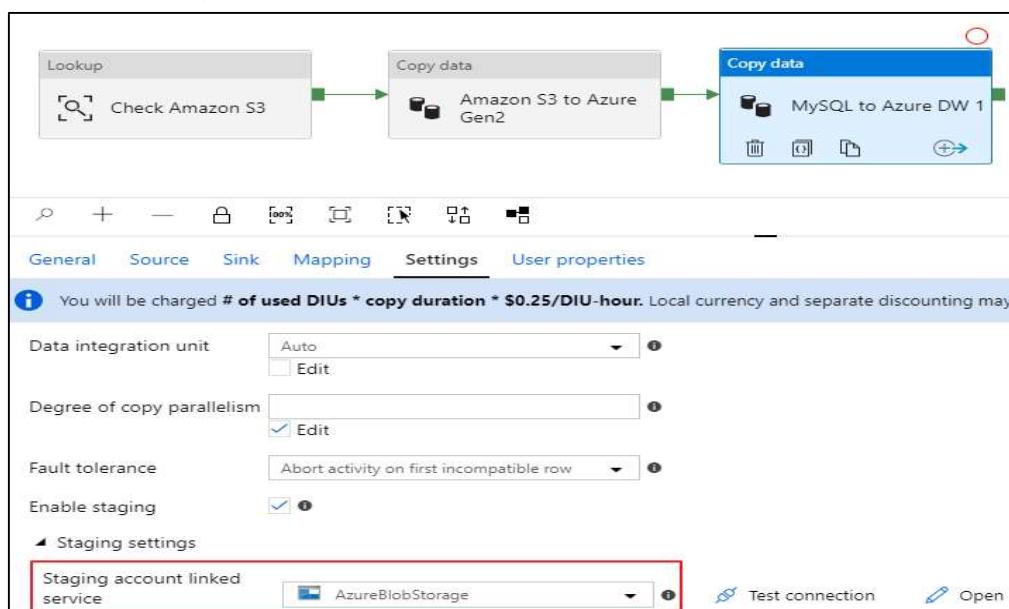
General	Source	Sink	Mapping	Settings	User properties																		
<p>You will be charged # of used DIUs * copy duration * \$0.25/DIU-hour. Local currency and</p> <table> <tr> <td>Data integration unit</td> <td>Auto</td> <td><input type="checkbox"/> Edit</td> </tr> <tr> <td>Degree of copy parallelism</td> <td></td> <td><input checked="" type="checkbox"/> Edit</td> </tr> <tr> <td>Fault tolerance</td> <td>Abort activity on first incompatible row</td> <td><input type="checkbox"/> Edit</td> </tr> <tr> <td>Enable staging</td> <td><input checked="" type="checkbox"/> <small>Add dynamic content [Alt+P]</small></td> <td><input type="button" value="New"/></td> </tr> <tr> <td colspan="3">► Staging settings</td> </tr> <tr> <td>Staging account linked service</td> <td>Select...</td> <td><input type="button" value="New"/></td> </tr> </table>						Data integration unit	Auto	<input type="checkbox"/> Edit	Degree of copy parallelism		<input checked="" type="checkbox"/> Edit	Fault tolerance	Abort activity on first incompatible row	<input type="checkbox"/> Edit	Enable staging	<input checked="" type="checkbox"/> <small>Add dynamic content [Alt+P]</small>	<input type="button" value="New"/>	► Staging settings			Staging account linked service	Select...	<input type="button" value="New"/>
Data integration unit	Auto	<input type="checkbox"/> Edit																					
Degree of copy parallelism		<input checked="" type="checkbox"/> Edit																					
Fault tolerance	Abort activity on first incompatible row	<input type="checkbox"/> Edit																					
Enable staging	<input checked="" type="checkbox"/> <small>Add dynamic content [Alt+P]</small>	<input type="button" value="New"/>																					
► Staging settings																							
Staging account linked service	Select...	<input type="button" value="New"/>																					

14. 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: **asostorageXXXX (스토리지 계정 명)**

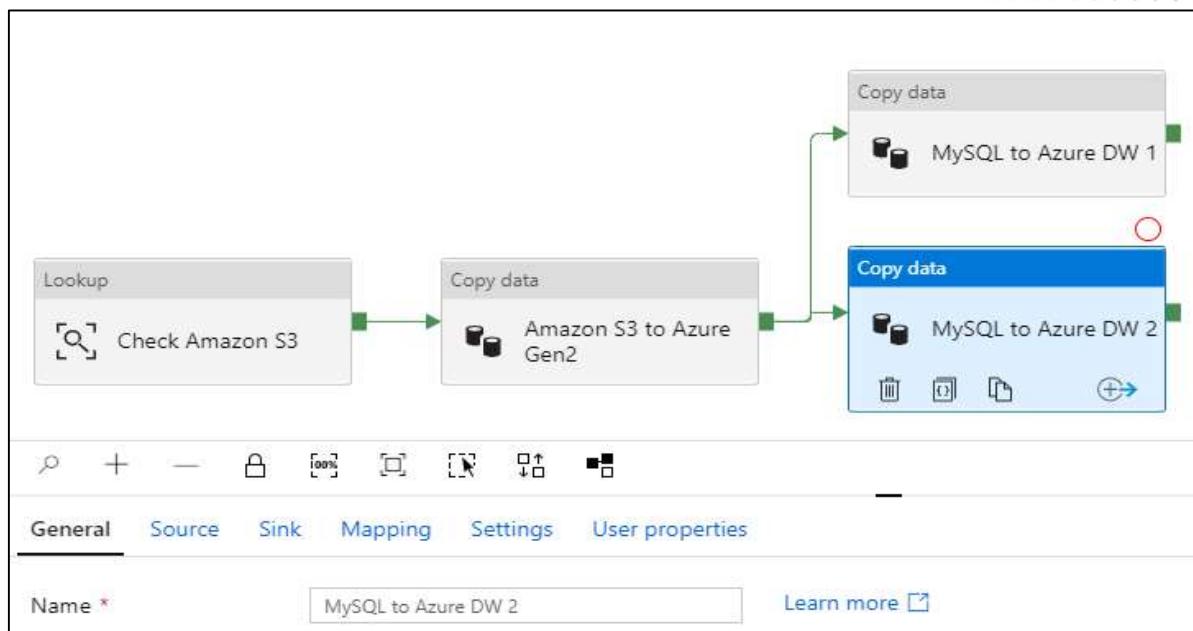


15. Settings 탭으로 이동하여 Enable staging 항목에 체크 후 Staging account linked service 항목을 **AzureBlobStorage**로 선택합니다.



16. MySQL to Azure DW 1 를 복사 후 붙여넣기 후 하단 General 탭에서 다음을 입력합니다.

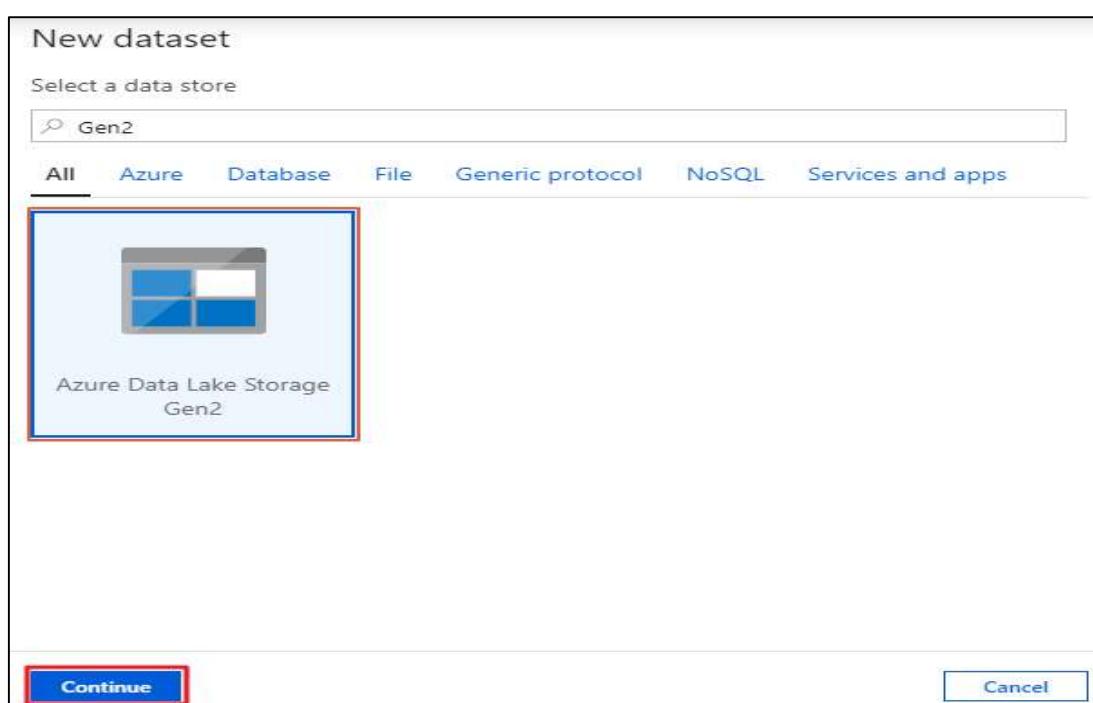
- ✓ Name: MySQL to Azure DW 2



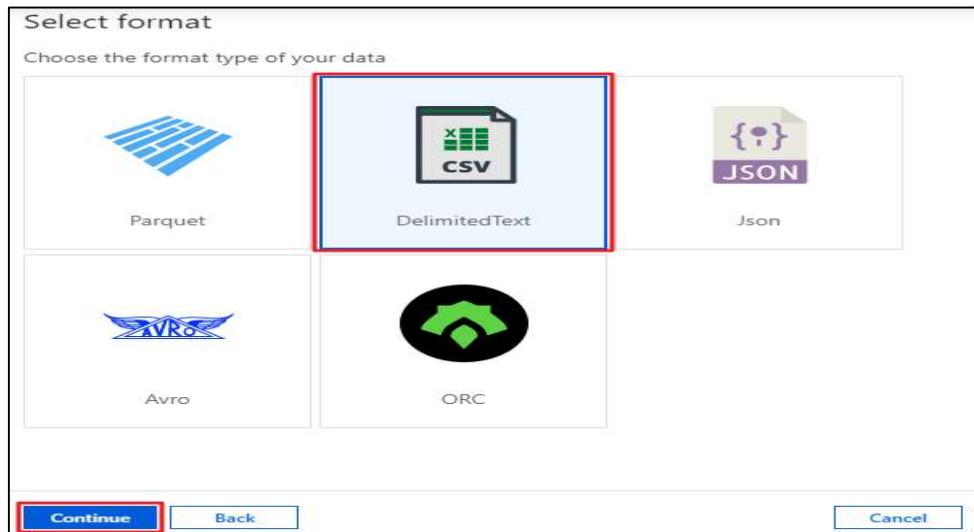
17. Source 탭으로 이동하여 Source dataset의 +New 를 클릭합니다.



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



19. format은 DelimitedText를 선택합니다



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

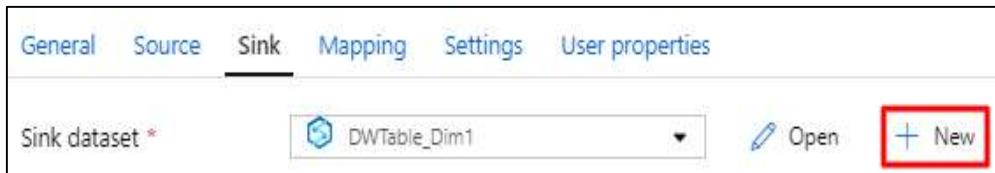
- ✓ Name: Gen2_Storage_Dim2
- ✓ Linked service : AzureDataLakeStorage
- ✓ File path : demo / mysql / DimGeography.csv (Browse 버튼 클릭하여 경로지정)
- ✓ First row as header : Check

The screenshot shows the 'Set properties' blade. It includes fields for Name (Gen2_Storage_Dim2), Linked service (AzureDataLakeStorage), File path (demo / mysql / DimGeography.csv, with the 'Browse' button highlighted with a red border), First row as header (checked), and Import schema (From connection/store selected). At the bottom are 'OK', 'Back', and 'Cancel' buttons.

21. Source탭의 Source Dataset이 세팅된 것을 확인합니다.

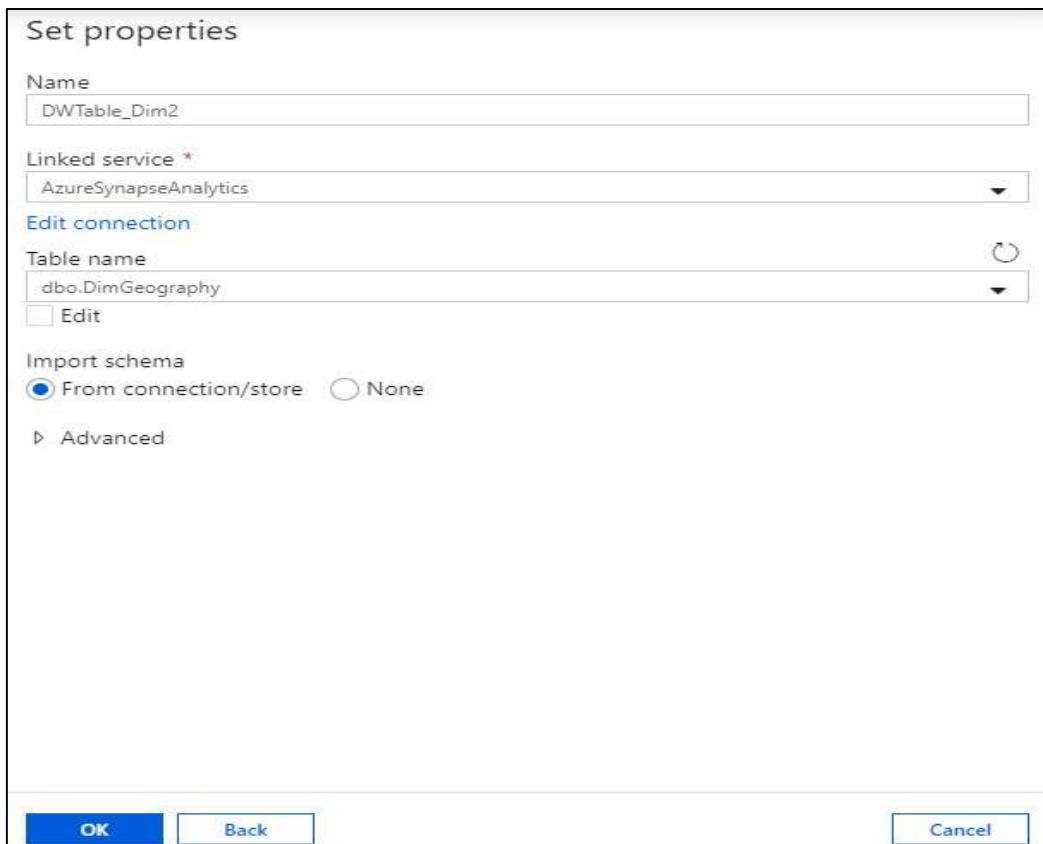


22. Sink탭으로 이동하여 Sink dataset의 +New를 클릭합니다.

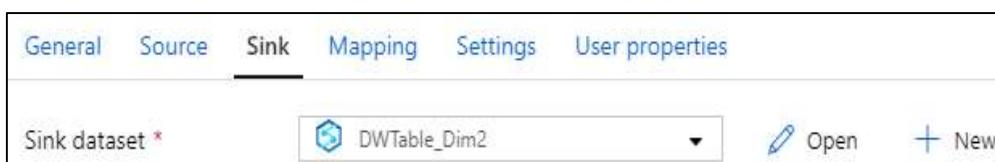


23. Set properties 화면이 열리면 아래와 같이 입력 후 OK 버튼을 클릭합니다.

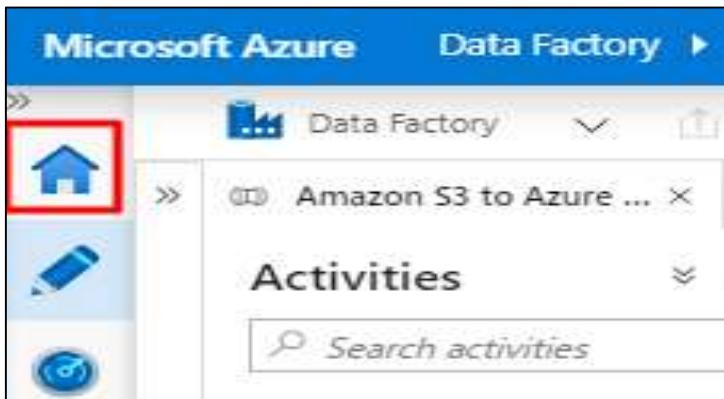
- ✓ Name: DWTable_Dim2
- ✓ Linked service : AzureSynapseAnalytics
- ✓ Table Name : dbo. DimGeography



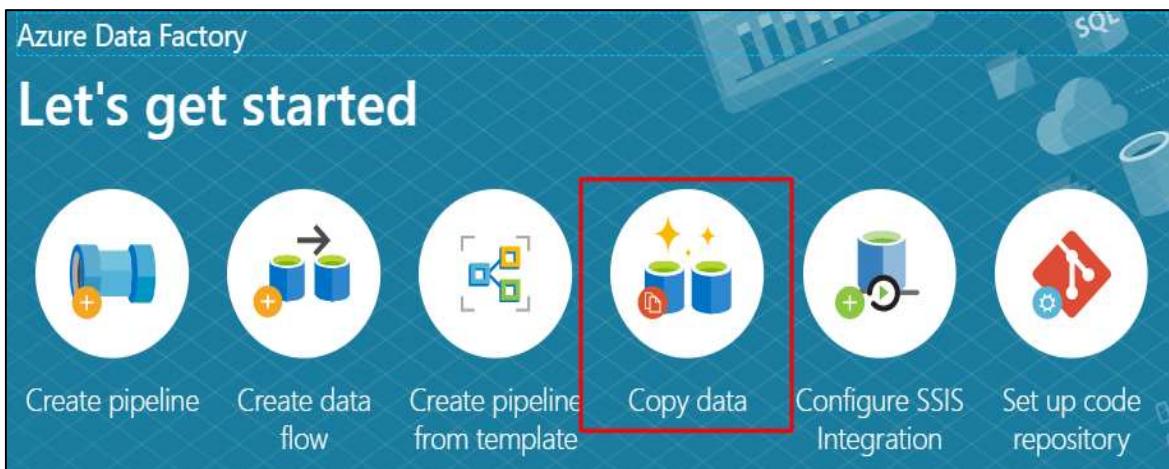
24. Copy data 액티비티의 Sink탭의 Sink Dataset이 세팅된 것을 확인합니다.



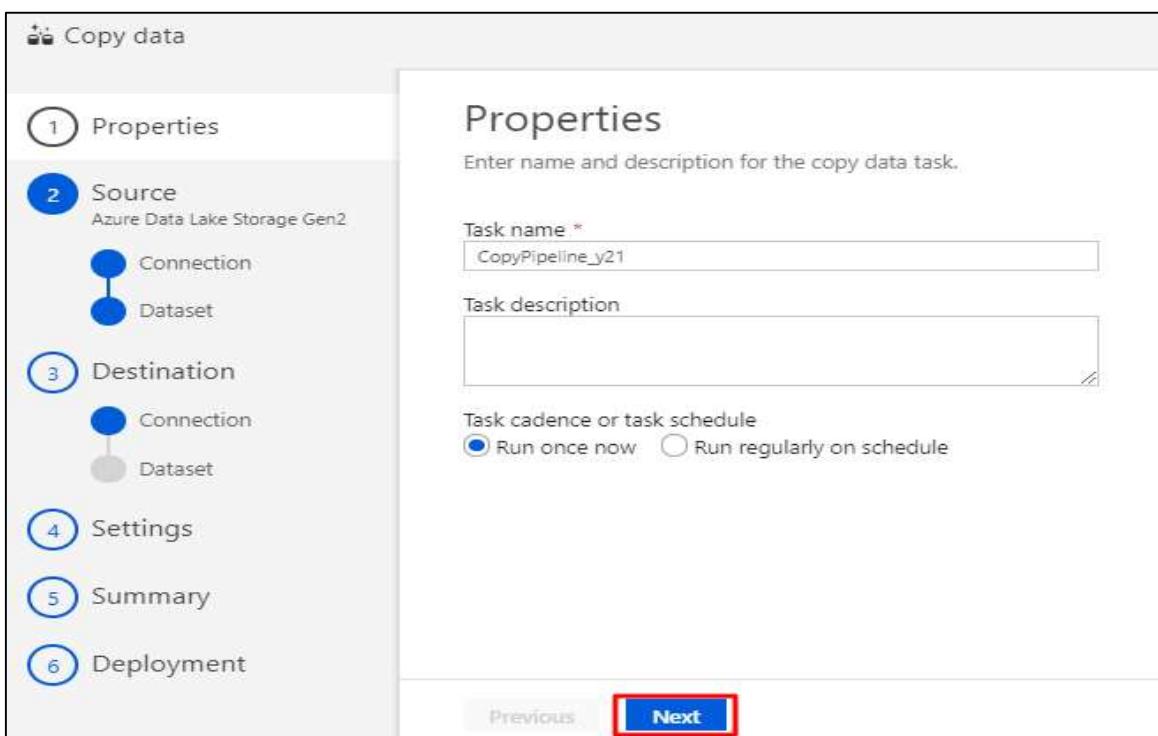
25. 왼쪽 메뉴에서 **Home** 아이콘을 클릭합니다.



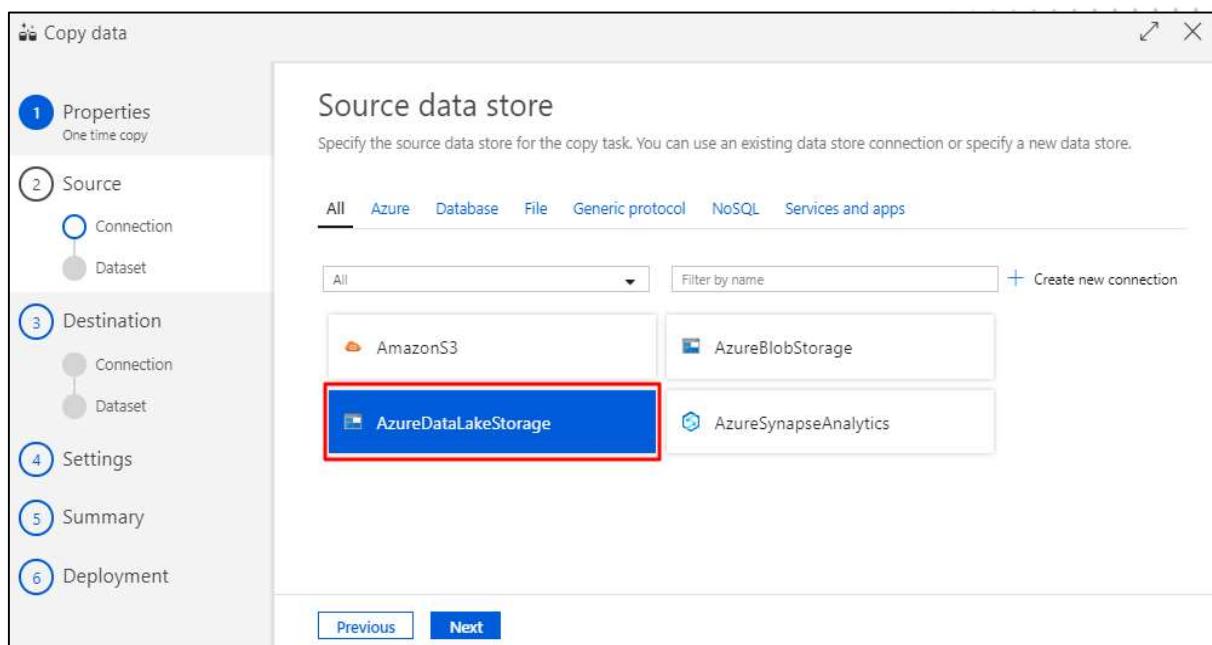
26. Data Factory 포털 홈 화면에서 **Copy Data**를 클릭합니다.



27. Properties 화면에서 **Next** 버튼을 클릭합니다.

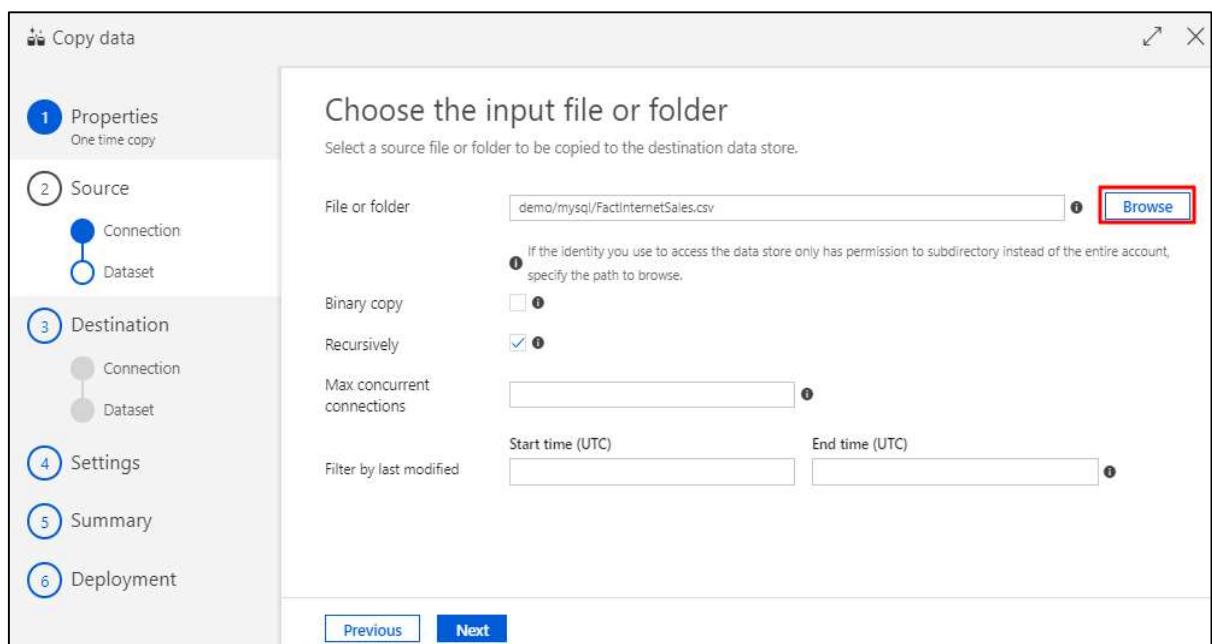


28. 기존에 만들어 놓았던 Gen2 스토리지인 **AzureDataLakeStorage**를 선택 후 Next 버튼을 클릭합니다.



29. Browse 버튼을 클릭하여 다음과 같이 경로를 지정합니다.

- ✓ File path : **demo / mysql / FactInternetSales.csv**



30. 파일 포맷 및 미리보기를 확인 후 Next 버튼을 클릭합니다.

File format settings

File format: Text format Detect text format

Column delimiter: Comma (,) Edit

Row delimiter: Auto detect (\r\n, \n, or \r\n\r\n)

Preview **Schema**

CarrierTrackingNumber	CurrencyKey	CustomerKey	CustomerPONumber	DiscountAmount	DueDateKey	ExtendedAmc
19	21768		0	20010713	\$3578.27	
39	28389		0	20010713	\$3399.99	

Previous **Next**

31. Destination data store 화면에서 **AzureSynapseAnalytics**를 선택 후 Next 버튼을 클릭합니다.

Destination data store

Specify the destination data store for the copy task. You can use an existing data store connection or specify a new data store.

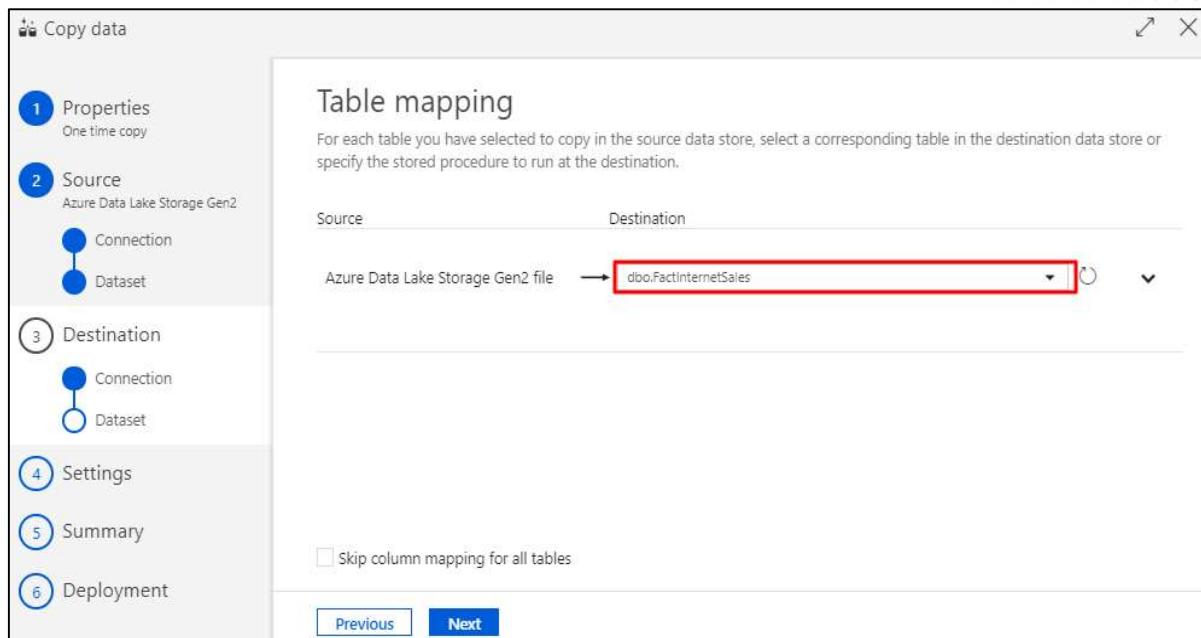
All **Azure** Database File Generic protocol NoSQL Services and apps

All Filter by name + Create new connection

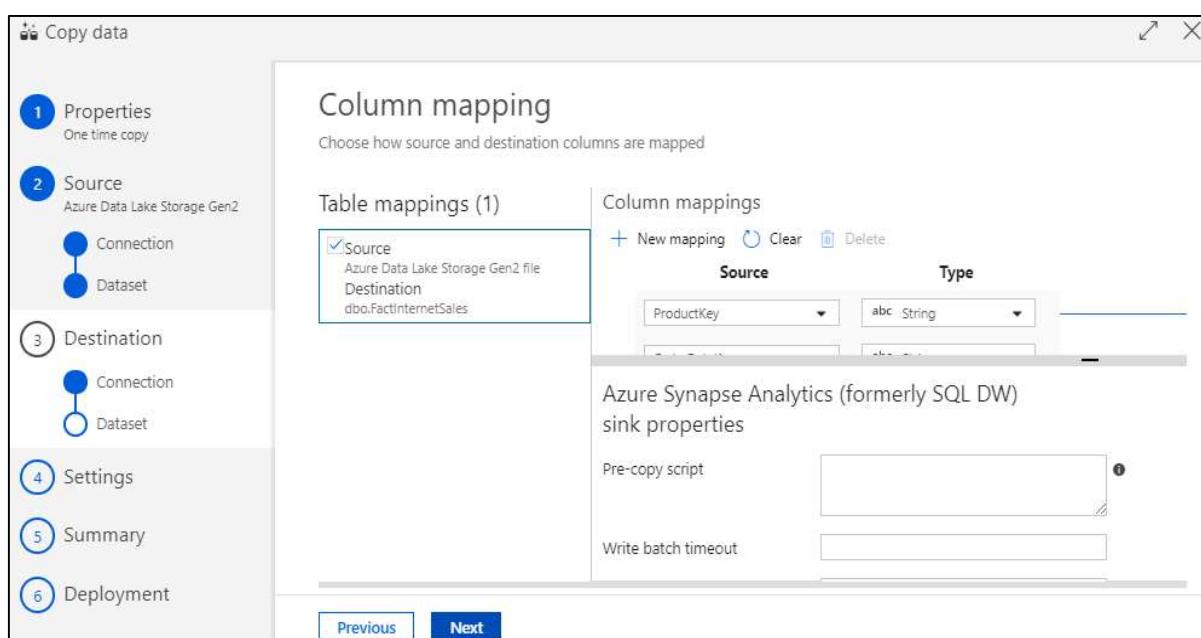
AzureSynapseAnalytics

Previous **Next**

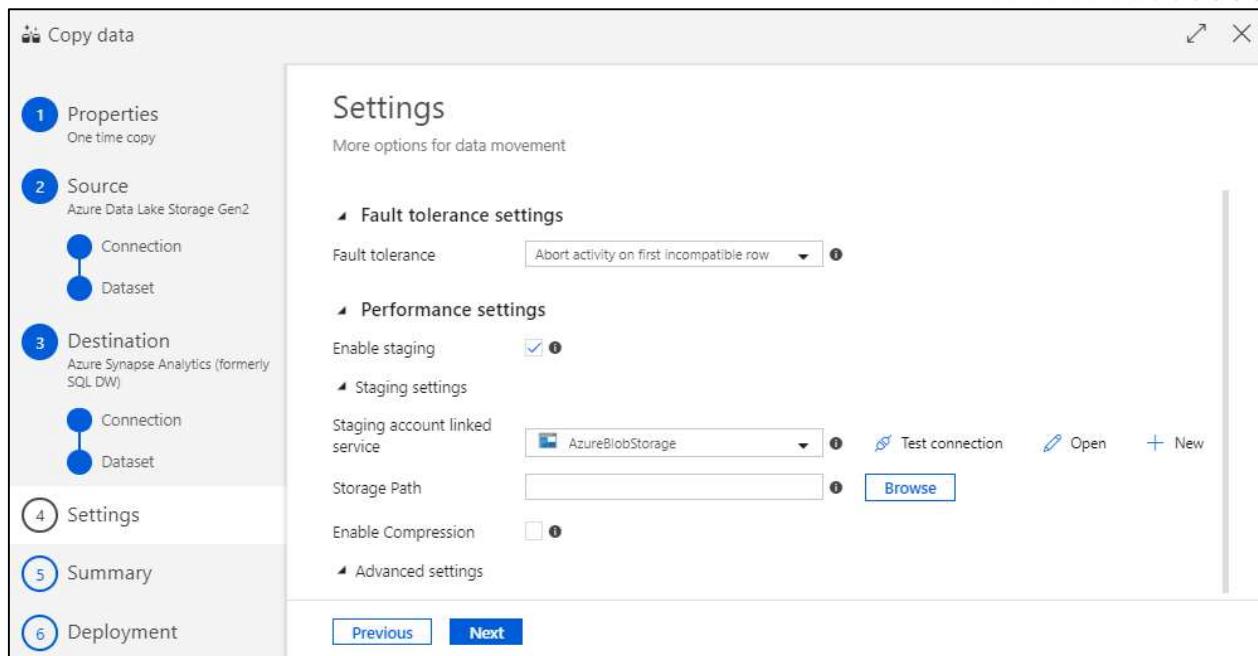
32. Destination 항목의 **dbo.factInternetSales**를 선택 후 Next 버튼을 클릭합니다.



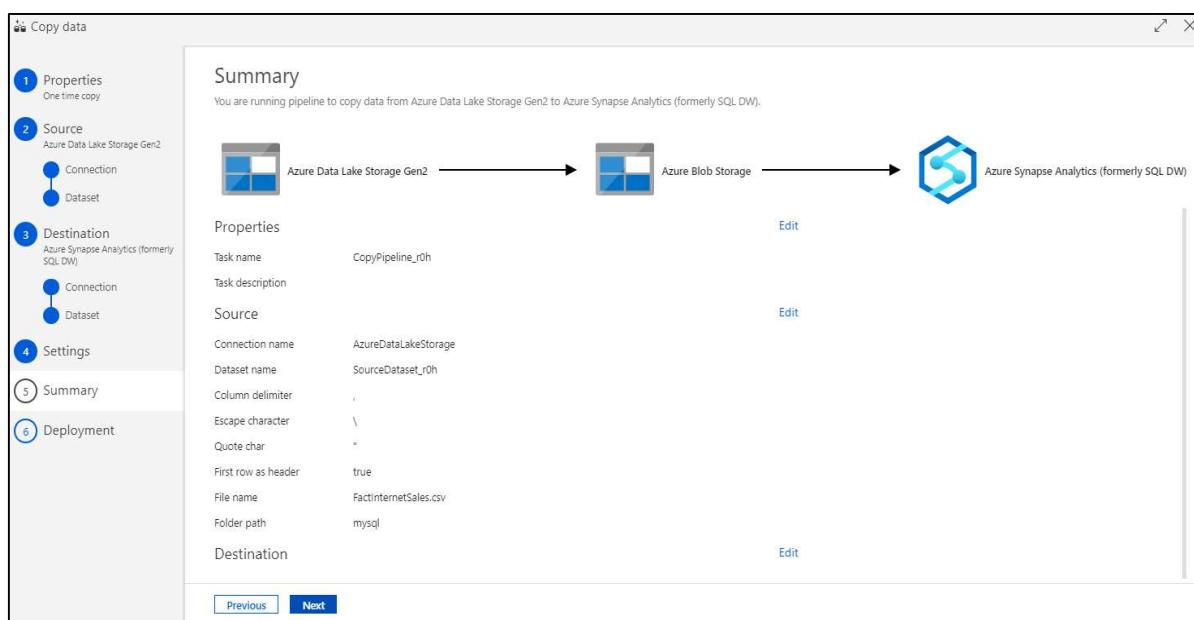
33. 컬럼 매핑 정보를 확인 후 Next 버튼을 클릭합니다.



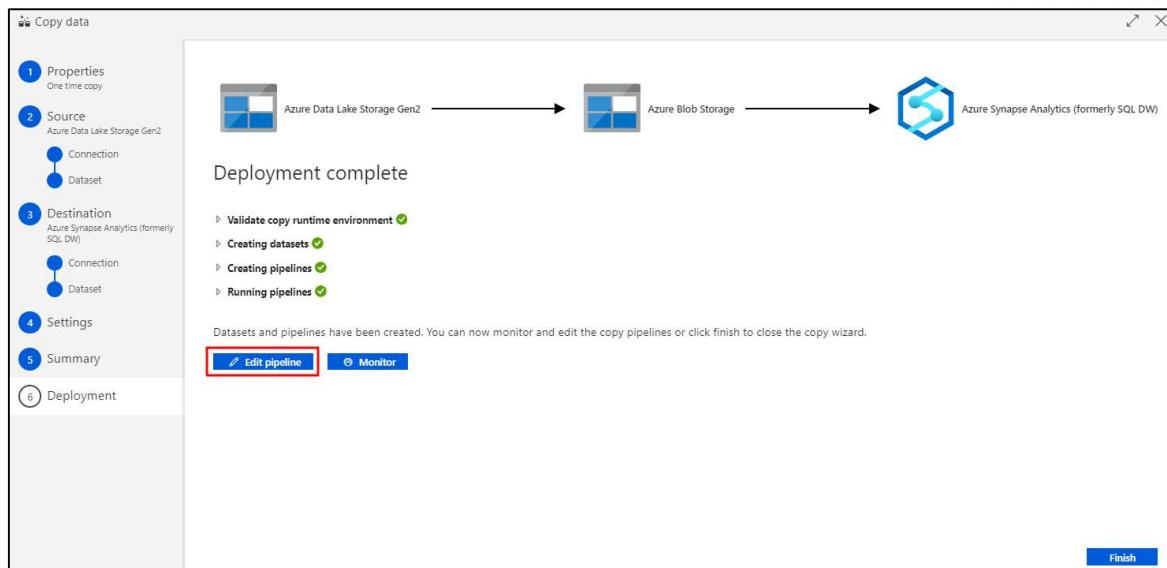
34. Settings 화면의 Enable staging 항목에 체크 후 Staging account linked service 항목의 **AzureBlobStorage**를 선택합니다.



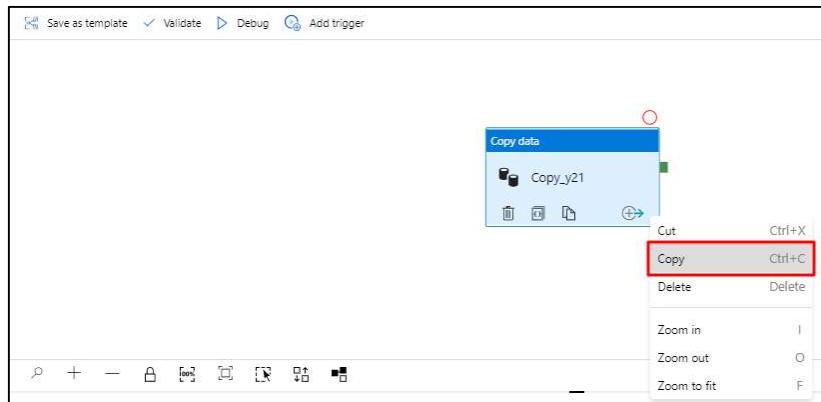
35. 요약된 내용을 확인 후 **Next** 버튼을 클릭합니다.



36. Deploy가 완료되면 **Edit pipeline** 버튼을 클릭하여 pipeline으로 이동합니다.

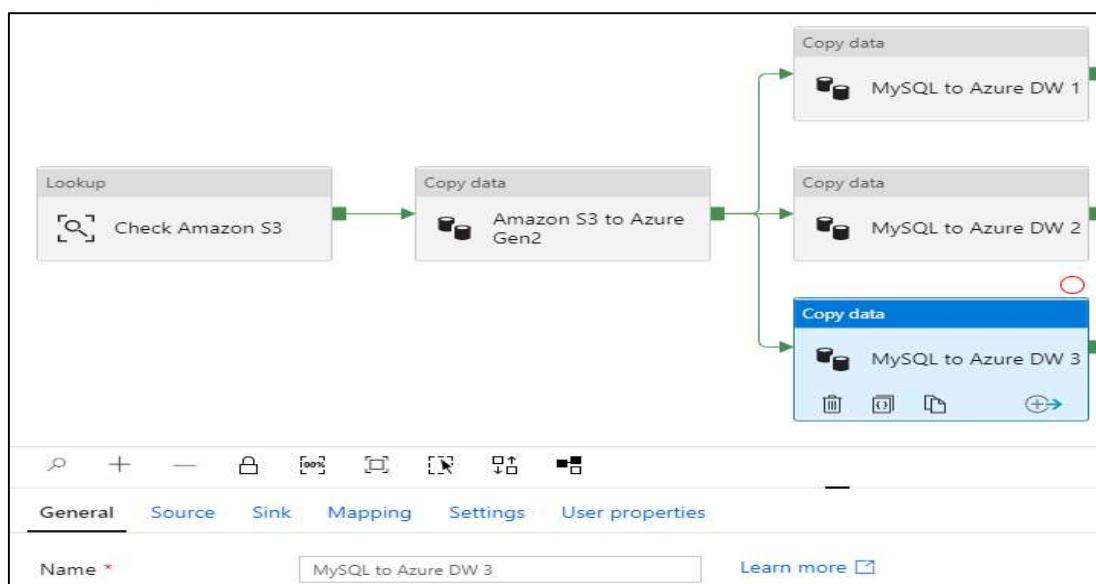


37. 생성된 copy data를 복사합니다.

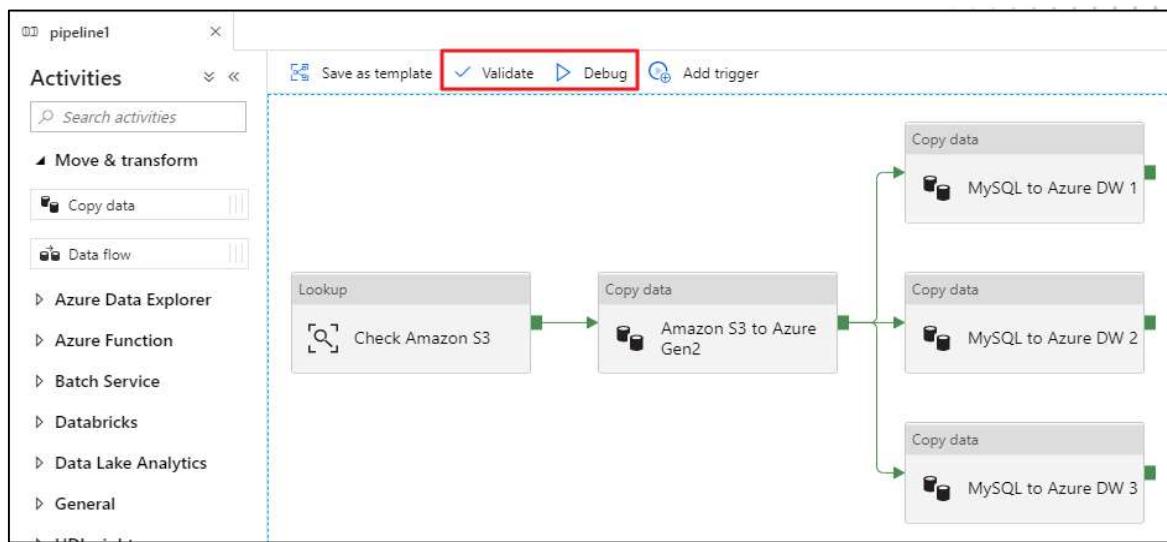


38. 기존에 작업중이던 파이프라인에 붙여넣기 한 후 Amazon S3 to Azure Gen2와 연결하여 General 탭에서 다음을 입력합니다.

- ✓ Name: MySQL to Azure DW 3



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



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

이름	타입
ASODatafactory0000	Data Factory(V2)
asodwserver0000	SQL Server
asostorage0000	스토리지 계정
CohoDW(asodwserver0000/CohoDW)	SQL Data Warehouse

2. 왼쪽 메뉴에서 쿠리 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]

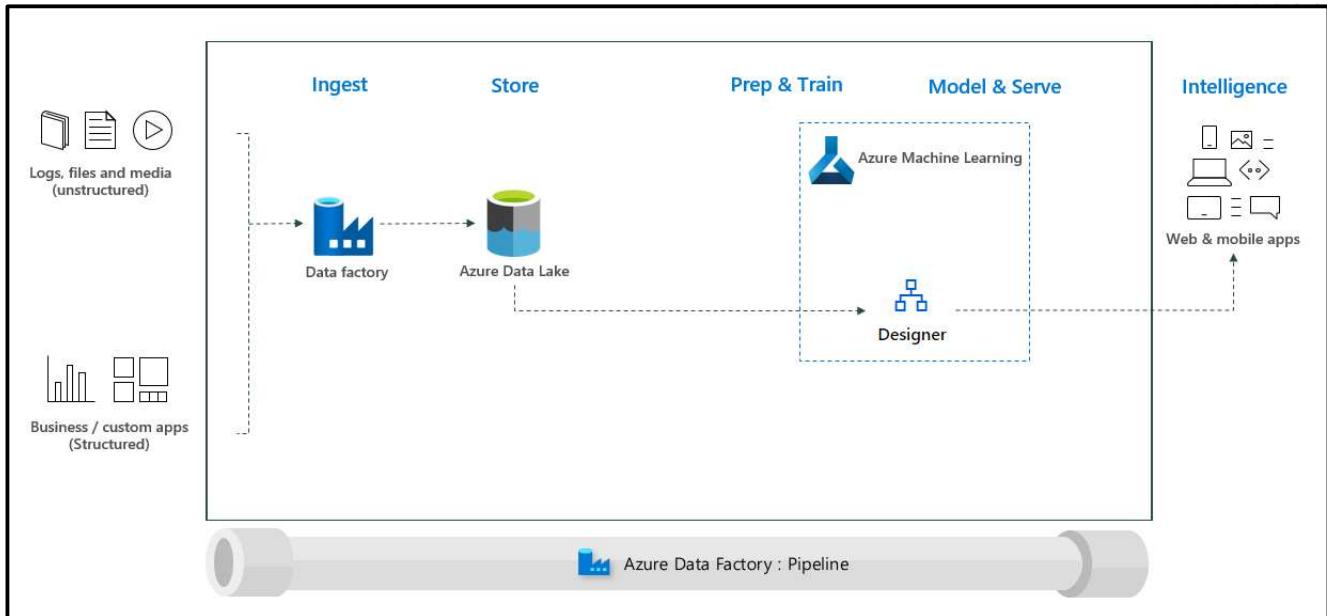
SELECT TOP(100) * FROM [dbo].[DimGeography]

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	

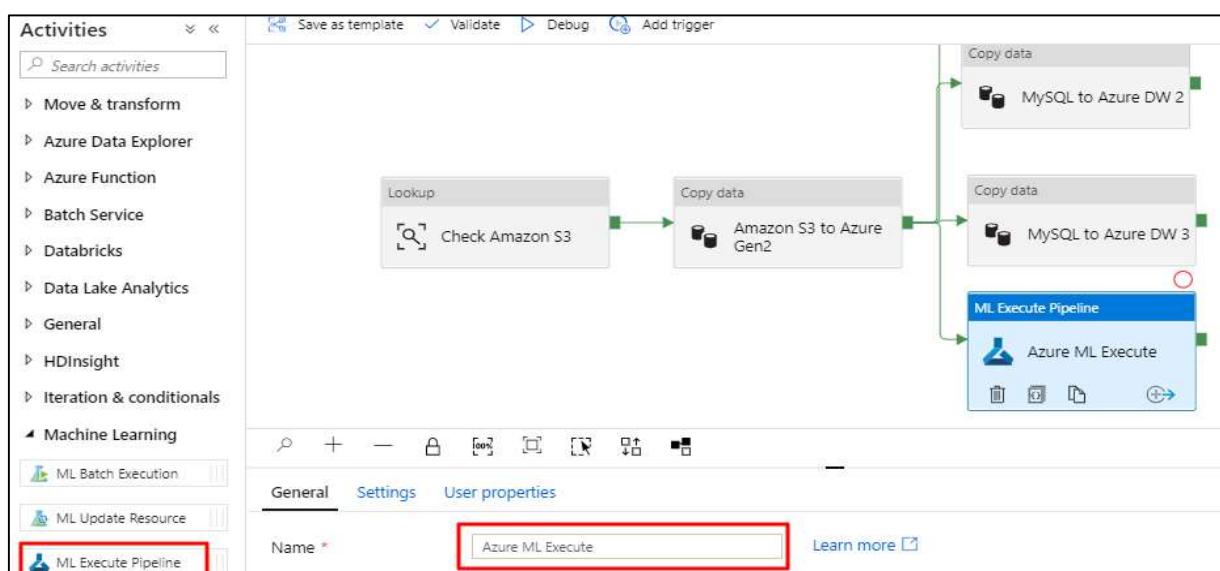
4.3. Task 3: Azure ML Execute

이 작업에서는 Azure ML 서비스에 연결하여 Task 1에서 복사한 파일을 이용하여 Model을 학습하는 ML 파이프라인을 실행시킵니다.



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:

Annotations: + New

Advanced: Connection successful Test connection

Create

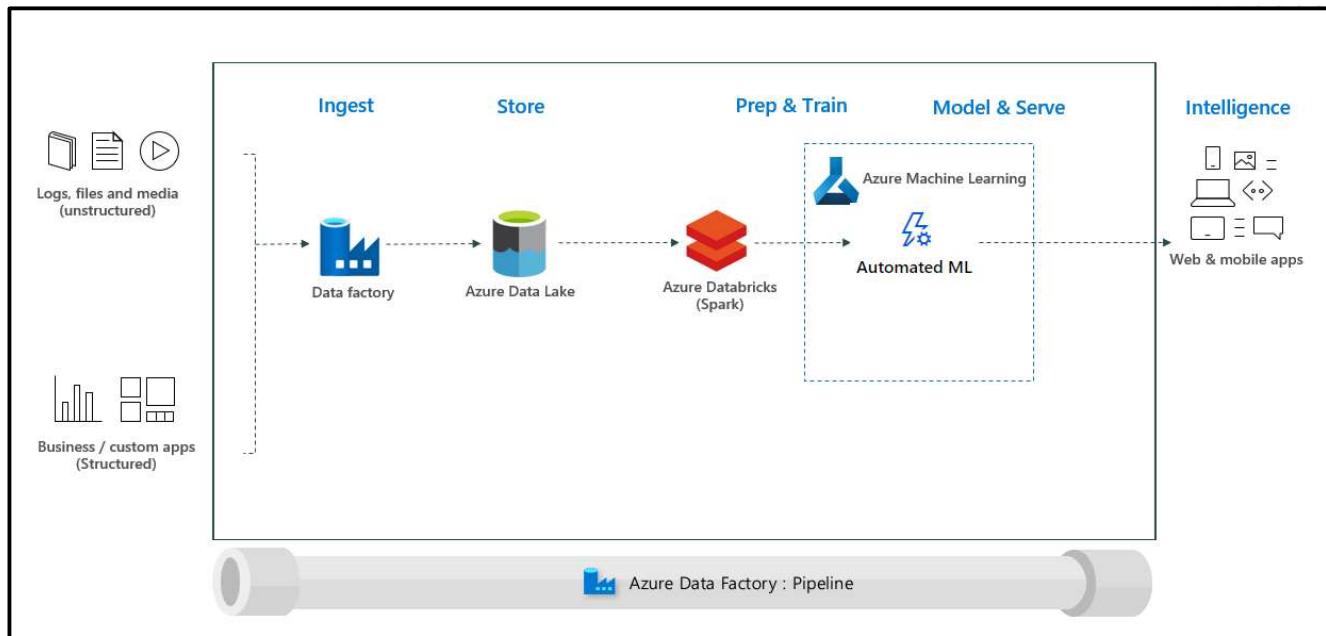
5. AML Service linked service 항목이 세팅된 것을 확인 후 아래의 정보를 입력합니다.

- ✓ 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"/>	<input type="button" value="Test connection"/> <input type="button" value="Open"/> <input type="button" value="New"/>
ML pipeline name	<input type="text" value="Logistic_Movie_Review_Sample 02-19-2..."/>	<input type="button" value="Open in Azure Portal"/>
ML pipeline ID *	<input type="text" value="cd916d3c-6de0-4e90-93a8-b8491d1bd..."/>	
Experiment name	<input type="text" value="Designer_test_v1"/>	

4.4. Task 4: PreProcessing and DataExport

이 작업은 Task 1에서 복사한 파일을 이용하여 Azure Databricks 서비스에 연결하여 Automated ML을 실행하기 위한 데이터셋을 만듭니다.



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

- ✓ Name: PreProcessing and DataExport

The screenshot shows the Azure Data Factory pipeline editor interface. On the left, the 'Activities' sidebar is open, showing various options like Move & transform, Azure Data Explorer, and Databricks. Under Databricks, 'Notebook' is selected and highlighted with a red box. The main workspace contains a sequence of activities connected by arrows: a 'Lookup' activity followed by two 'Copy data' activities ('Amazon S3 to Azure Gen2' and 'MySQL to Azure DW 2'). A third 'Copy data' activity ('MySQL to Azure DW 3') is shown below. A 'Notebook' activity labeled 'PreProcessing and DataExport' is also present. The pipeline is named 'PreProcessing and DataExport' in the General tab. The General tab is currently selected, while other tabs like Azure Databricks, Settings, and User properties are also visible.

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

The screenshot shows the 'General' tab of the Azure Databricks linked service settings page. A dropdown menu labeled 'Databricks linked service' is displayed, with a red box highlighting the '+ New' button next to it. Other tabs like 'Azure Databricks', 'Settings', and 'User properties' are also visible at the top.

3. 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**
- ✓ Select cluster : **Existing interactive cluster**
- ✓ Access token : **dapid1fb7b77a6d0e7f9fb2ccf337367a9a5**
- ✓ 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:
 New job cluster
 Existing interactive cluster
 Existing instance pool

Domain/Region *: https://koreacentral.azuredatabricks.net

Access token (selected)
Azure Key Vault

Choose from existing clusters *:

Annotations:

Connection successful

Create **Test connection** **Cancel**

4. Databricks linked service 항목이 세팅된 것을 확인합니다.

General Azure Databricks **Settings** ¹ User properties

Databricks linked service *: **AzureDatabricks**

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

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

General Azure Databricks **Settings** User properties

Notebook path *: /Users/kch8306@datasolutionmlg.onmicrosoft.com/PreProcessing_EDU **Browse** Open

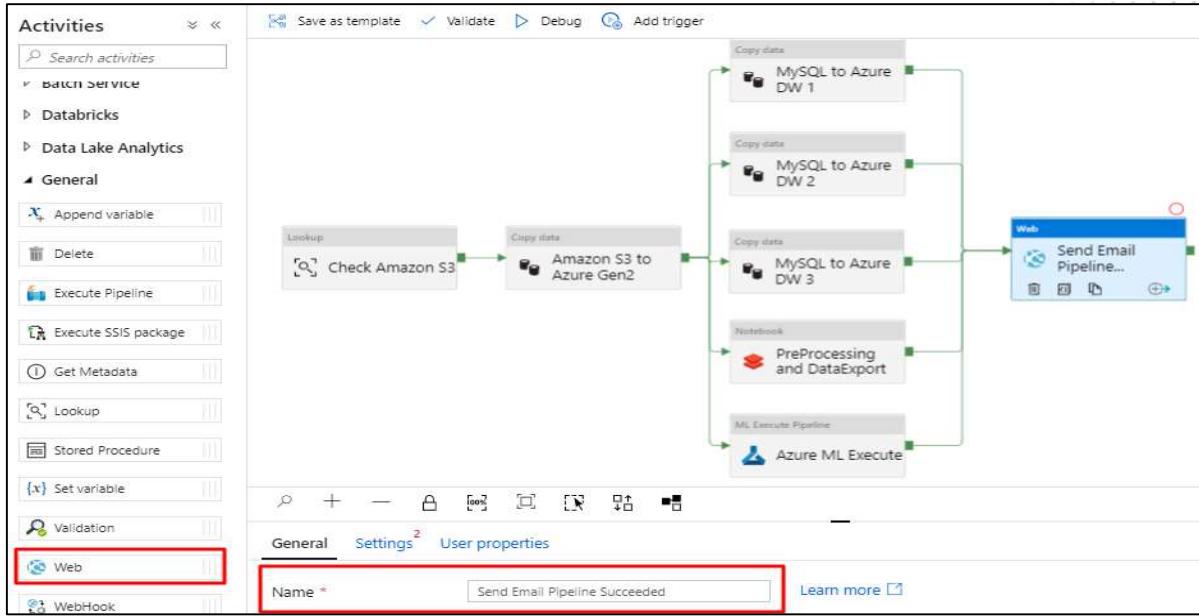
▶ Base parameters

▶ Append libraries

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

```
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

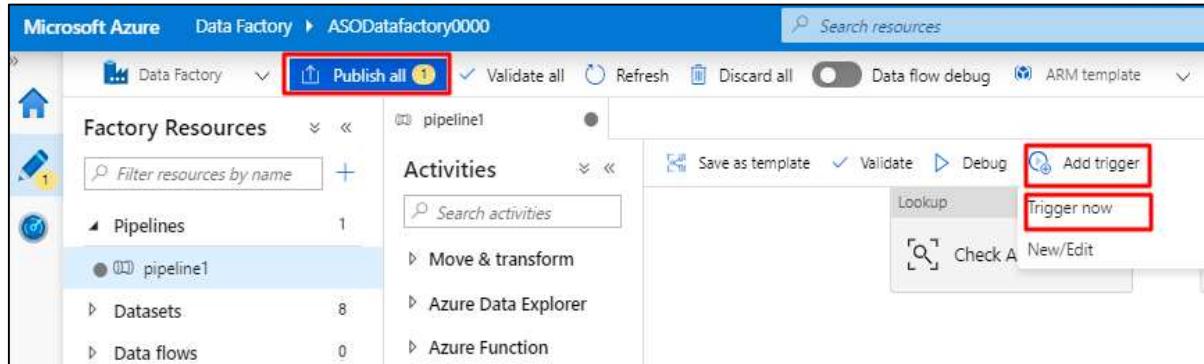
- ✓ Headers: +New 클릭 후 NAME 항목에 Content-Type, VALUE 항목에 application/json 추가

- ✓ Body:

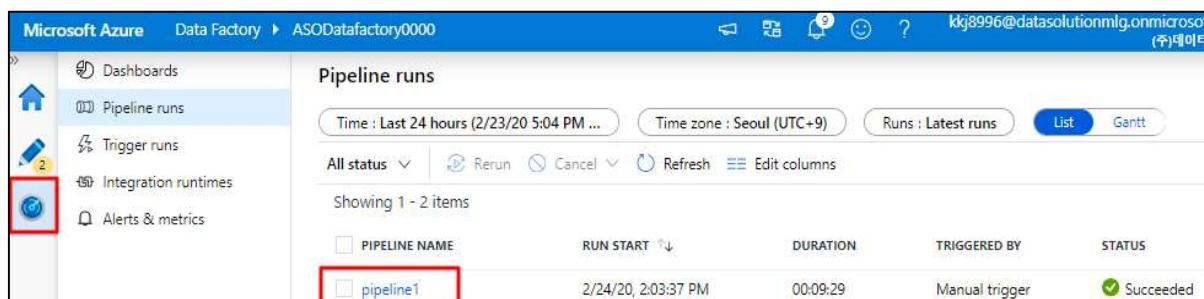
```
{"DataFactoryName": "@{pipeline().DataFactory}", "PipelineName": "@{pipeline().Pipeline}", "Subject": "Pipeline Succeeded", "Message": "Succeeded!", "EmailTo": "본인의 메일계정"}
```

General	Settings	User properties						
URL *	http://me2.do/xVJvqnxn							
Method *	POST							
Headers *	+ New <input type="button" value="Delete"/> <table border="1"> <tr> <td><input type="checkbox"/></td> <td>NAME</td> <td>VALUE</td> </tr> <tr> <td></td> <td>Content-Type</td> <td>application/json</td> </tr> </table>		<input type="checkbox"/>	NAME	VALUE		Content-Type	application/json
<input type="checkbox"/>	NAME	VALUE						
	Content-Type	application/json						
Body	{"DataFactoryName": "@{pipeline().DataFactory}", "PipelineName": "@{pipeline().Pipeline}", "Subject": "Pipeline Succeeded", "Message": "Succeeded!", "EmailTo": "본인의 메일계정"}							

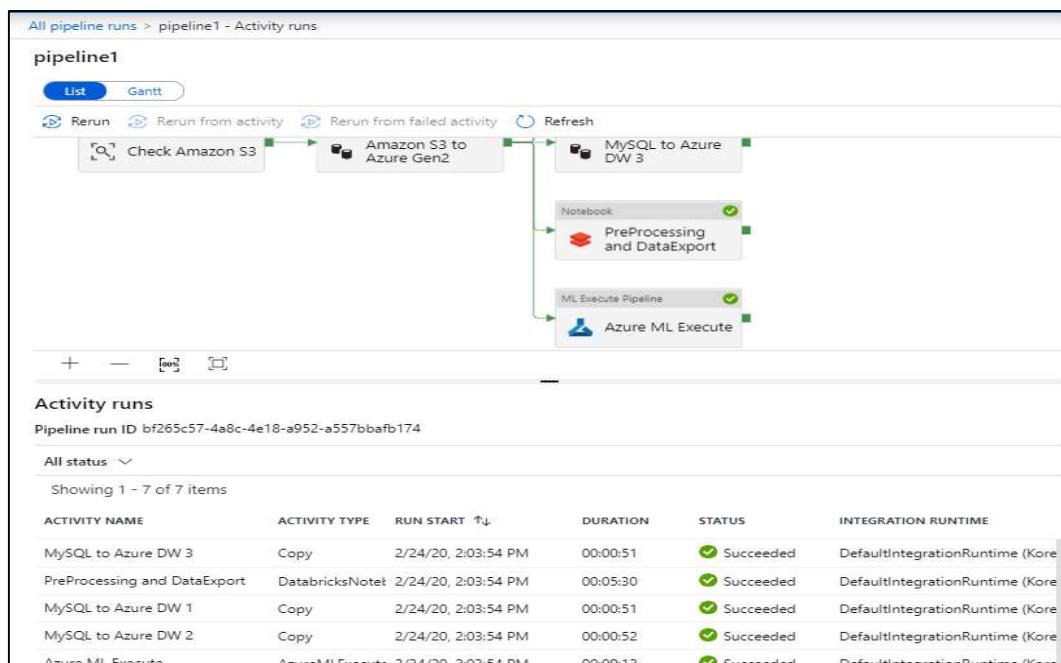
3. 최종 실행을 위해 왼쪽 상단의 **Publish All** 버튼 클릭 후 **Add trigger** 의 Trigger now를 클릭합니다.



4. 왼쪽 메뉴의 Monitor 아이콘을 클릭하여 진행 상태를 확인합니다.



5. PIPELINE NAME을 클릭하여 상세화면으로 이동합니다.

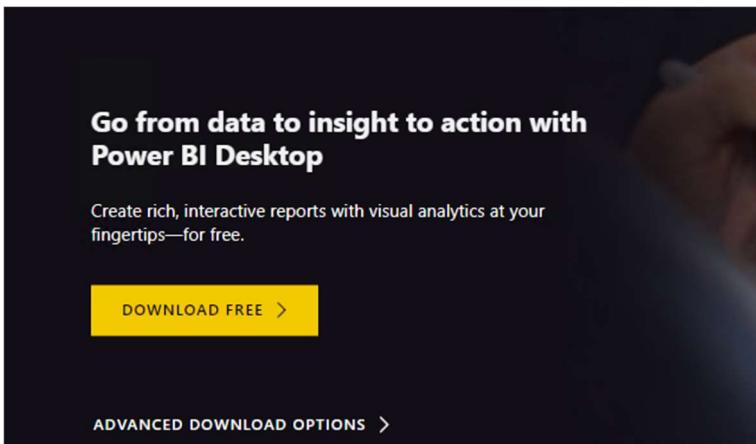


5. Exercise 3: Visualize data with Power BI Desktop

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.



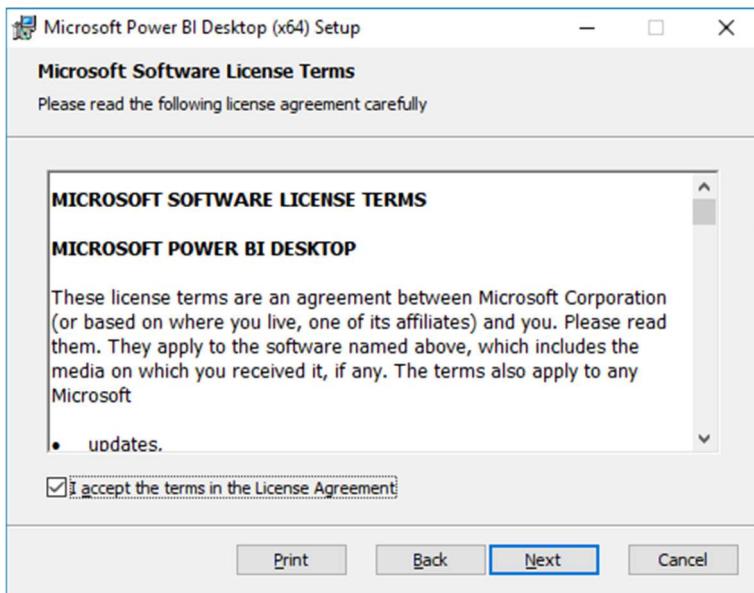
3. 설치파일을 실행합니다.



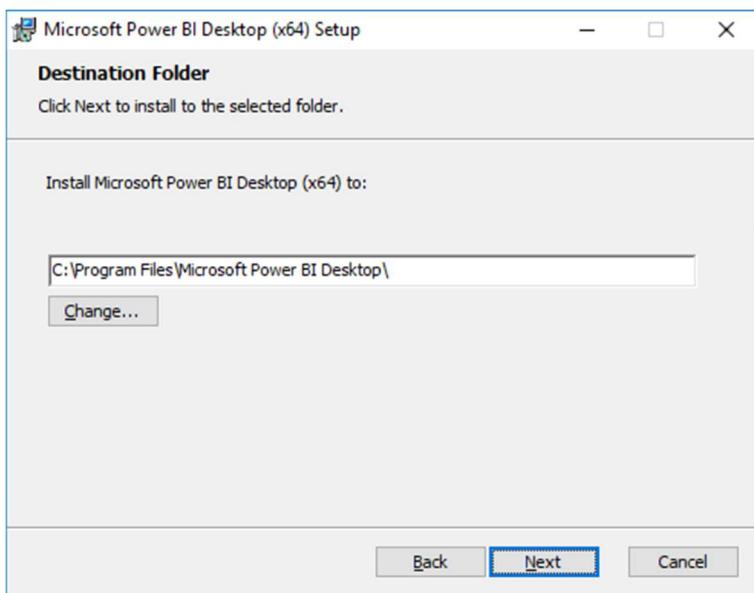
4. Welcome 화면에서 다음을 클릭합니다.



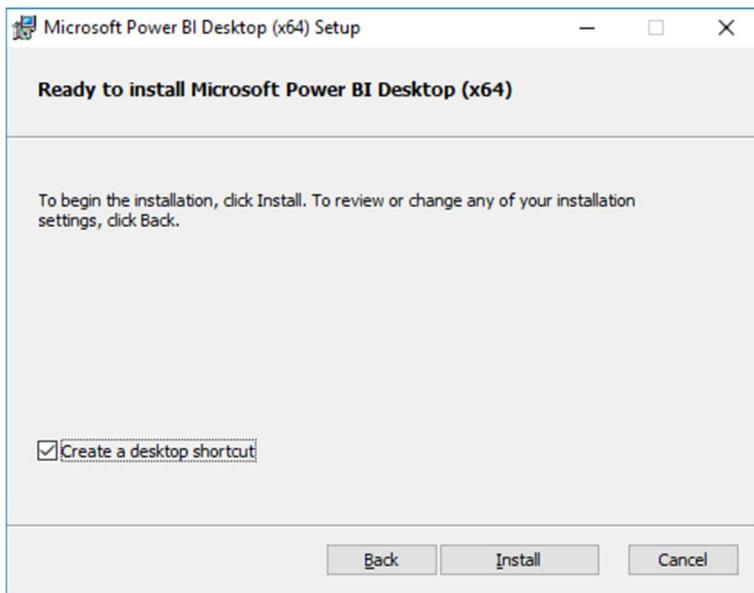
5. 사용권 계약에 동의하고 다음을 선택합니다.



6. 지정 폴더를 디폴트로 두고 다음을 선택합니다.



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

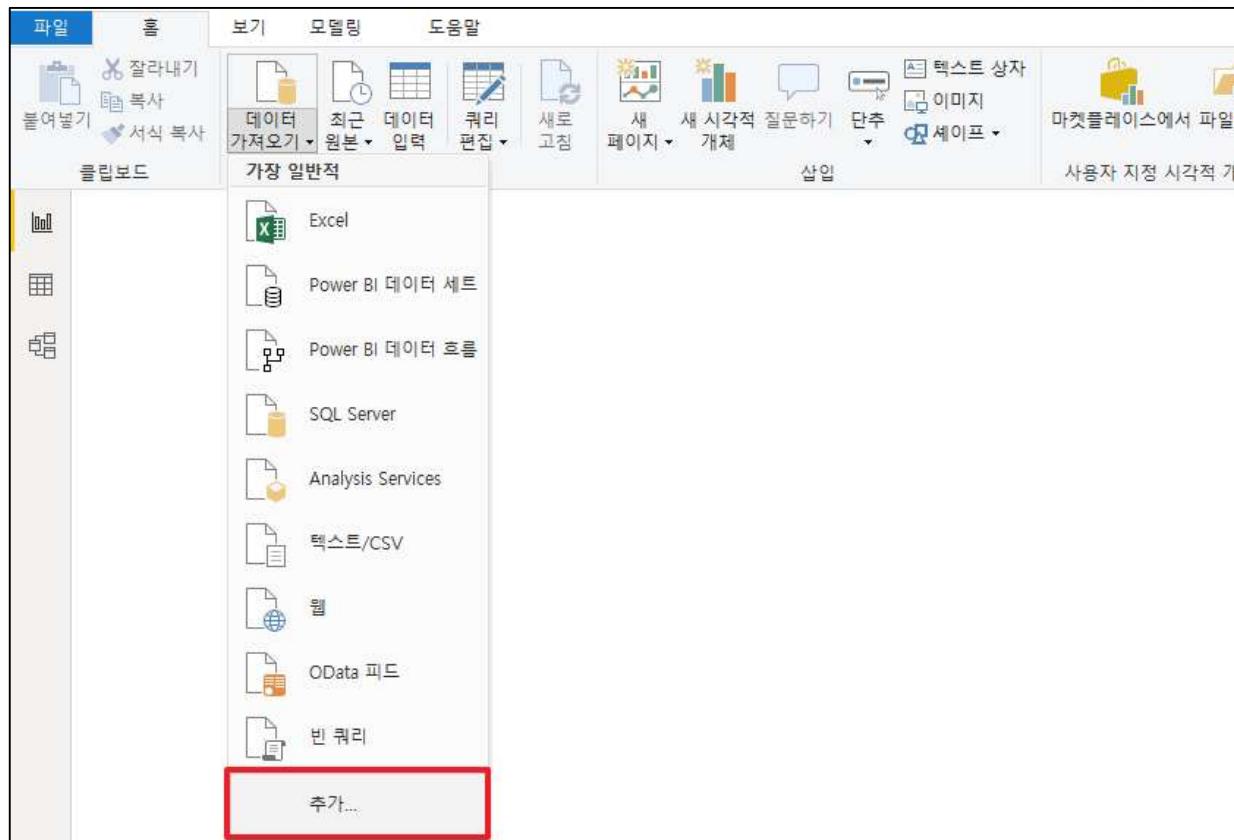


8. Microsoft Power BI Desktop 시작이 선택되어 있는지 확인하고 마침을 선택합니다.

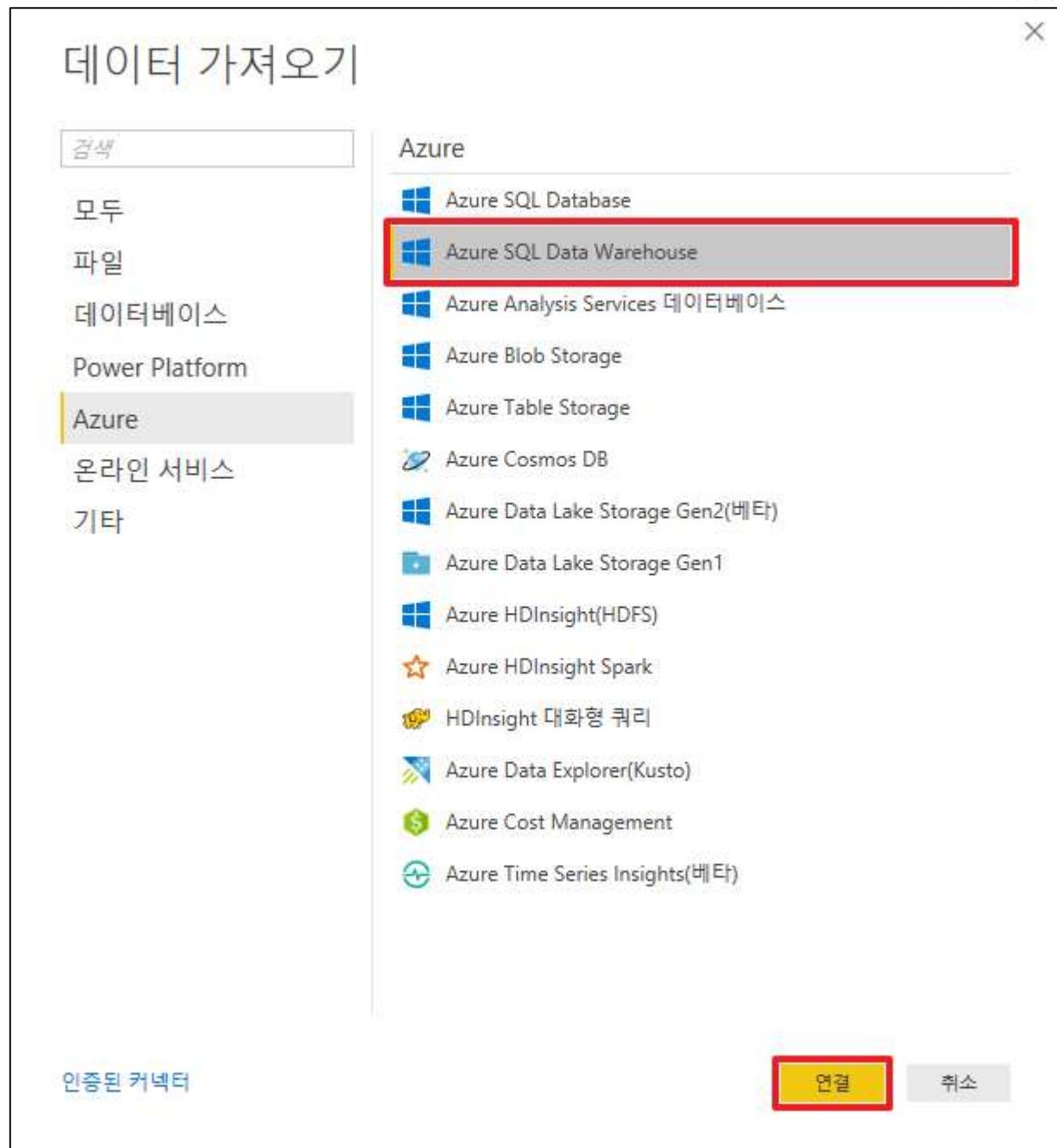


5.2. Task 2: Query data with Power BI Desktop

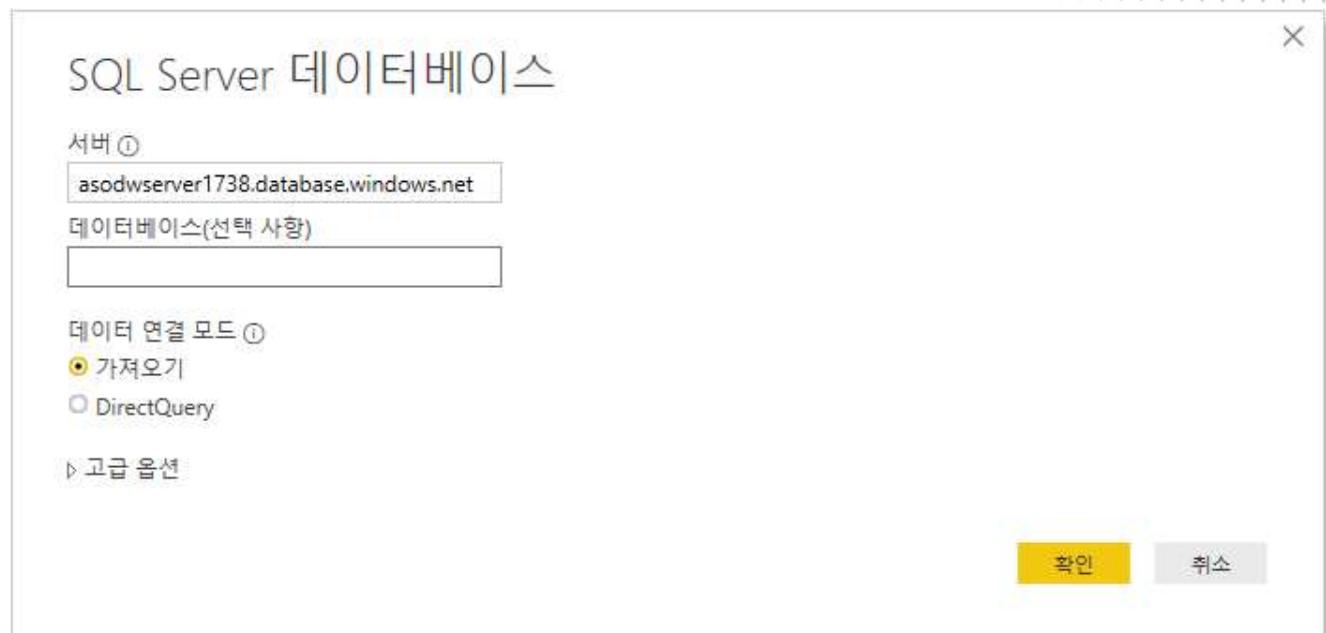
- Power BI에서 데이터 가져오기 버튼을 클릭한 후 추가...을 선택합니다.



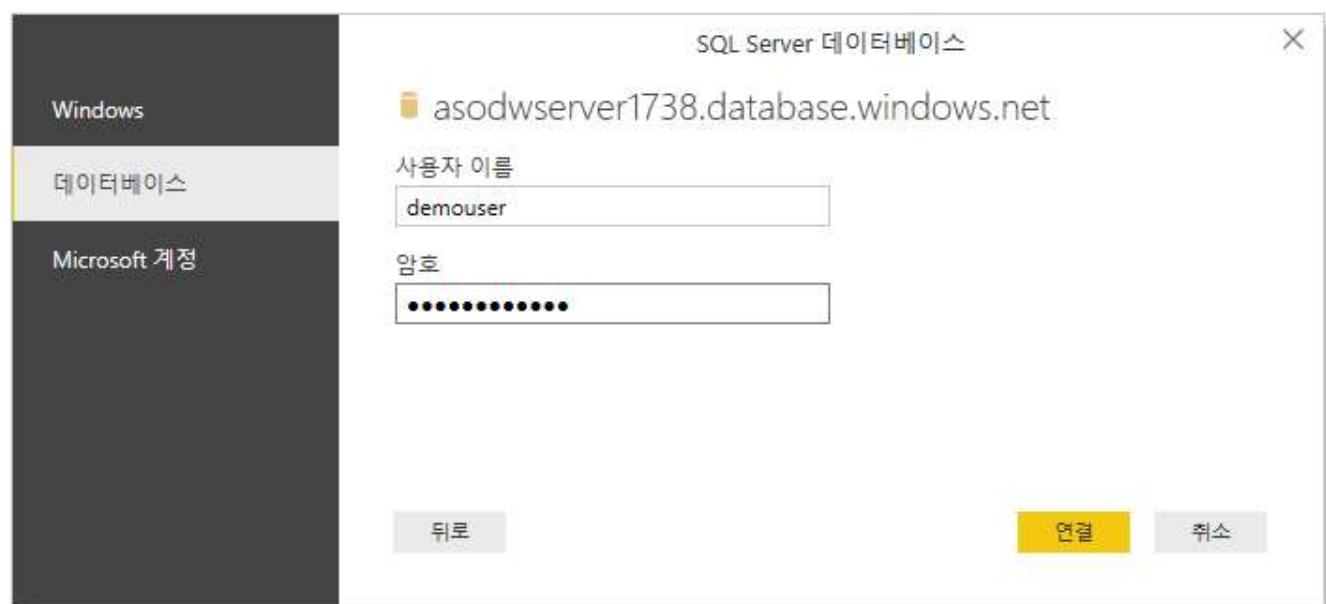
- 데이터 가져오기 창에서 왼쪽 목록에서 Azure를 선택합니다. 그런 다음 Azure SQL Data Warehouse를 선택하고 연결을 클릭합니다.



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

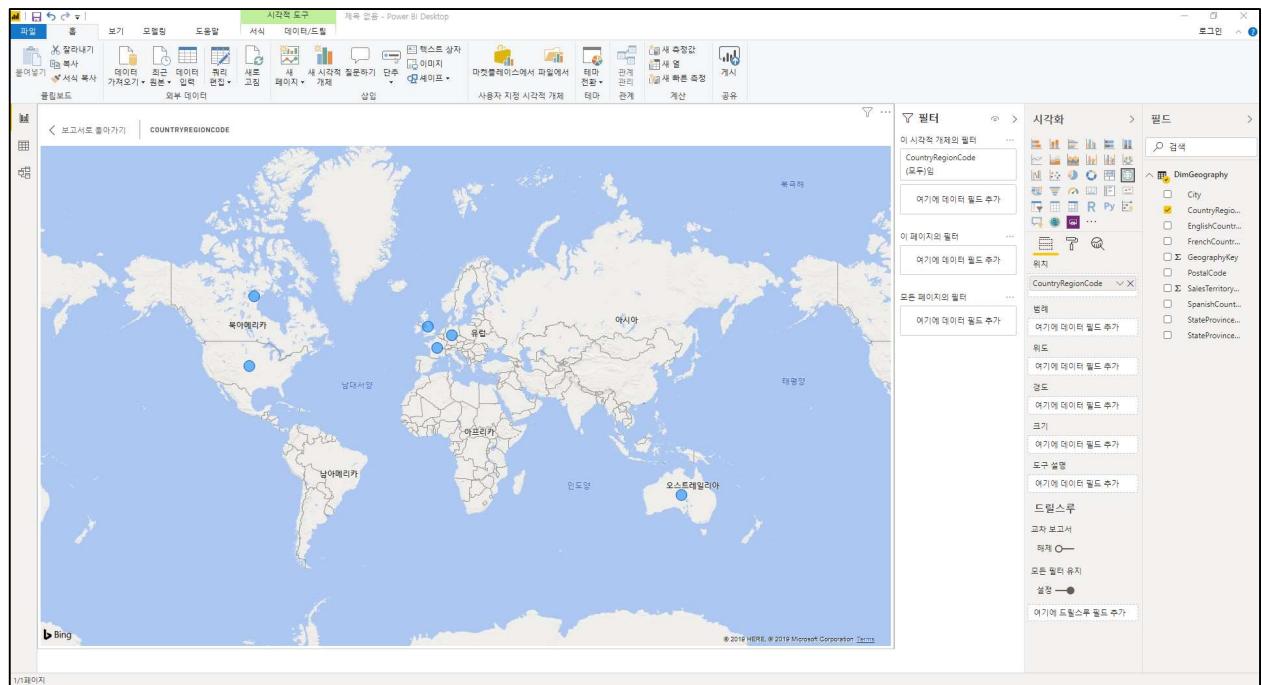
FactInternetSales

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	

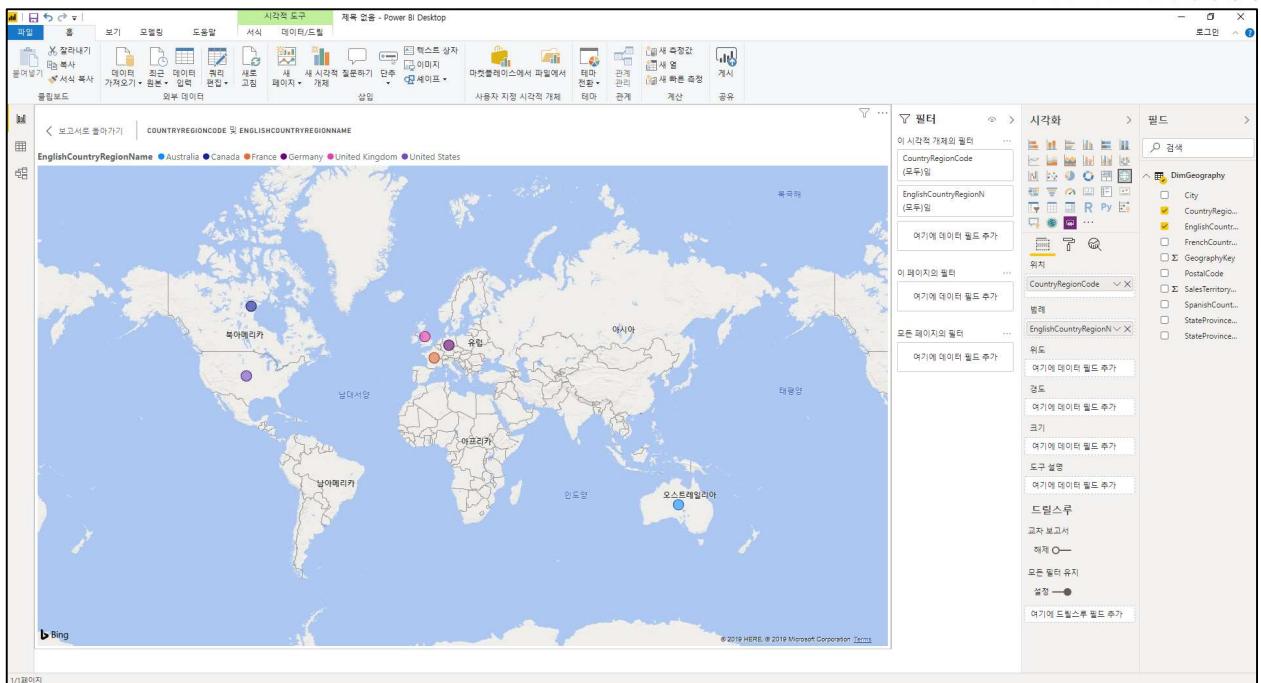
관련 테이블 선택

로드 데이터 변환 취소

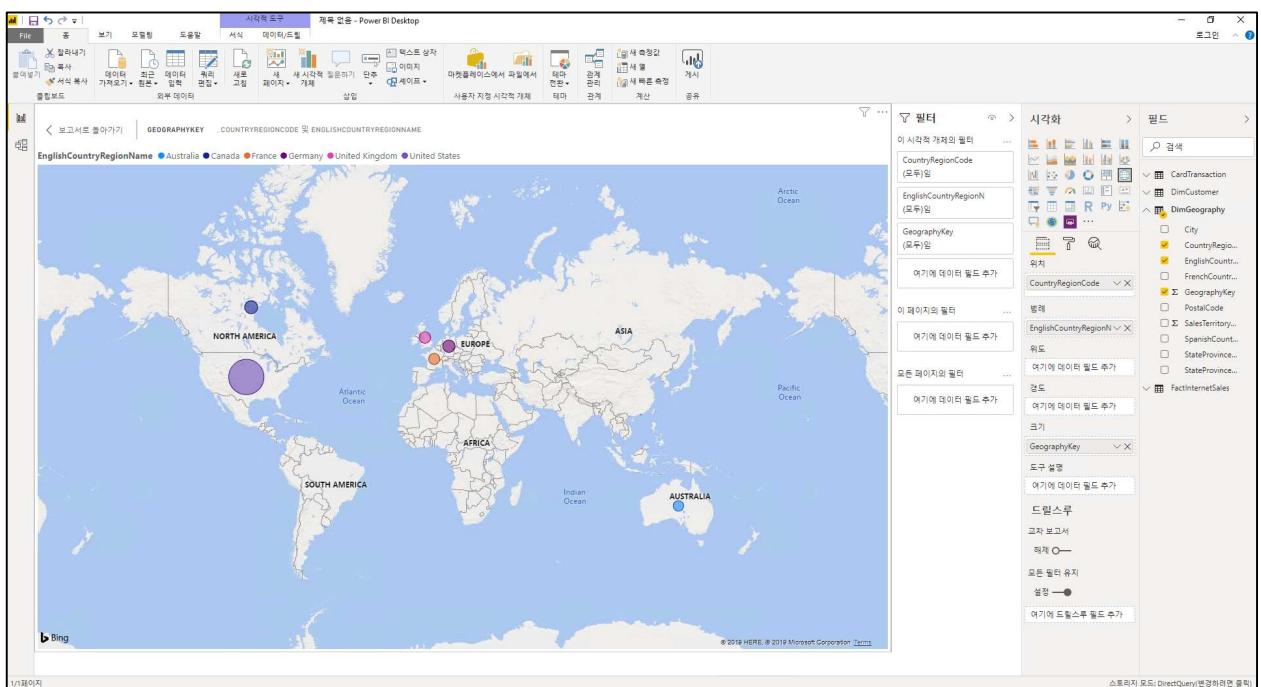
6. 오른쪽에 있는 필드 블레이드에서 DimGeography 테이블을 확장하고 CountryRegionCode 옆에 있는 상자를 선택합니다. Power BI는 지리 데이터를 인식하고 Map 시각화가 자동으로 시작됩니다.



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



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



실습이 완료되었습니다. 이번 실습을 통해 Azure Data Factory를 이용하여 전체 데이터의 흐름을 관리하는 과정에 대해 알아보았습니다.

- ✓ Azure Data Factory를 이용한 각기 다른 위치의 Data 이동
- ✓ Azure Data Factory에서 Azure에서 제공되고 있는 서비스 연결

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