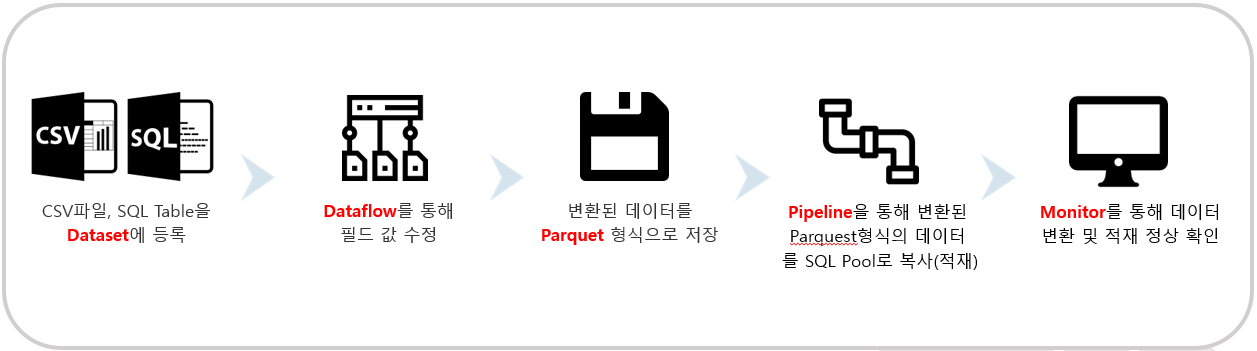
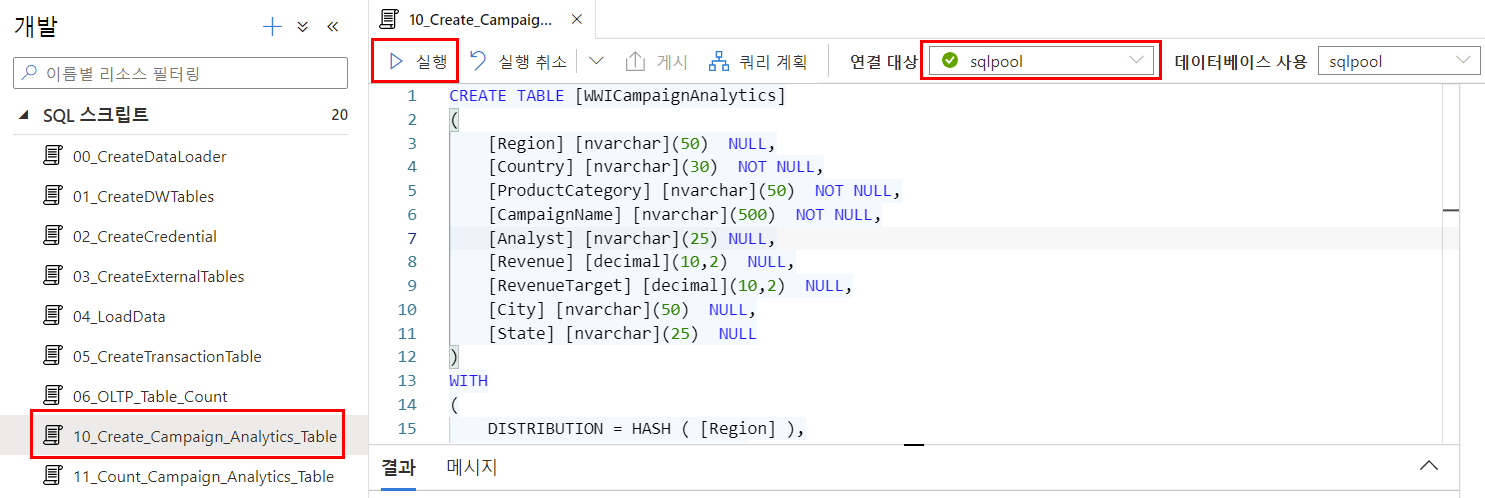
**작업 n – Create & Populate Campaign Analytics Table**

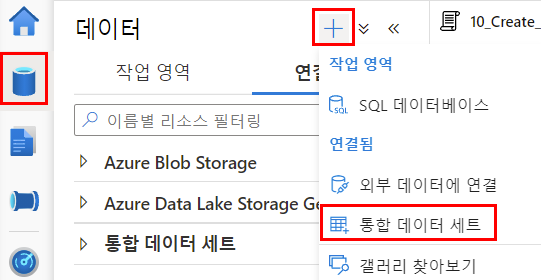
1. 시나리오



1. 왼쪽 메뉴의 “개발” 메뉴를 선택하고 **Create\_Campaign\_Analytics\_Table**   
   스크립트를 열어 실행합니다.

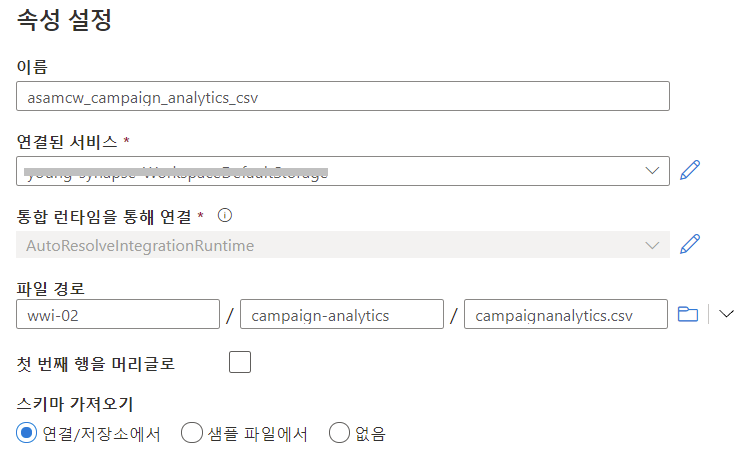


1. “데이터” 메뉴로 이동하여 + 버튼을 누르고 “통합 데이터 세트”를 생성합니다.



1. **“새 통합 데이터 세트”**에 Azure Data Lake Storage Gen2를 선택, **“형식 선택”**에 DelimitedText를 선택합니다.
2. 아래와 같이 속성 설정을 입력하고 확인 버튼을 누릅니다.

|  |  |
| --- | --- |
| Field | Value |
| 이름 | Enter **asamcw\_campaign\_analytics\_csv** |
| 연결된 서비스 | Select **<ADLS Gen2 Account Name>**  각자 가지고 있는 스토리지 계정 선택 |
| 파일 경로 – Container | Enter **pipeline** |
| 파일 경로 – Directory | Enter **convert** |
| 파일 경로 – File | Enter **campaignAnalytics.csv** |
| 첫 번째 행을 머리글로 | Unchecked |
| 스키마 가져오기 | Select **연결/저장소에서** |



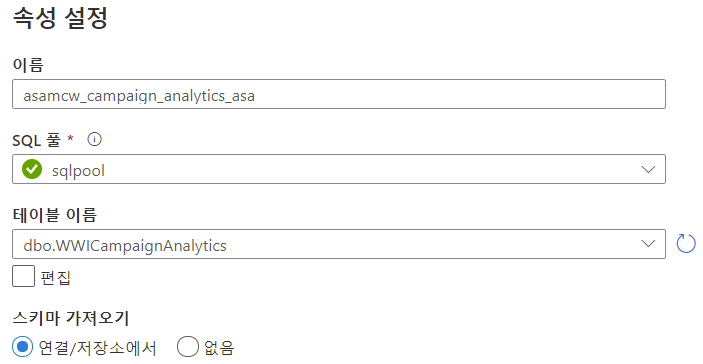
campaignAnalytics.csv

convert

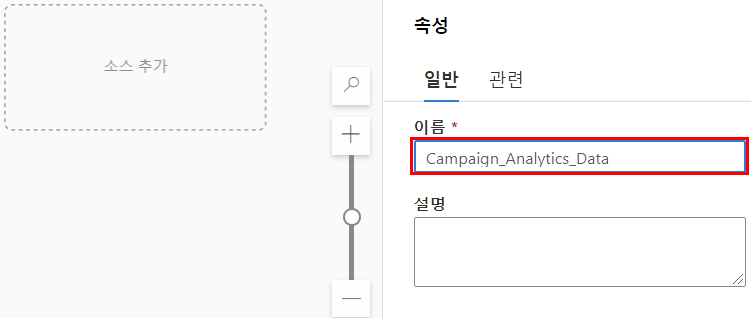
pipeline

1. 다시 + 를 누르고 **“통합 데이터 세트”**를 생성합니다.
2. “새 통합 데이터 세트”에 **Azure Synapse 전용 SQL pool**을 선택합니다.
3. 아래와 같이 속성 설정을 입력하고 확인 버튼을 누릅니다.

|  |  |
| --- | --- |
| Field | Value |
| 이름 | Enter **asamcw\_campaign\_analytics\_asa** |
| SQL 풀 | Select **sqlpool** |
| 테이블 이름 | Select **dbo.WWICampaignAnalytics** |
| 스키마 가져오기 | Select 연결/저장소에서 |

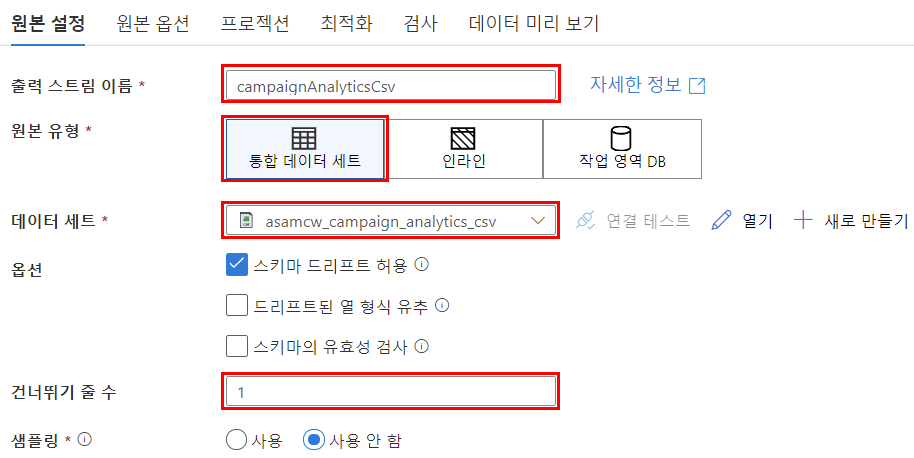


1. “모두 게시” 버튼을 누르고 저장합니다.
2. “개발” 메뉴로 돌아와서 +를 누르고 “**데이터 흐름**”을 생성합니다.
3. 속성값에 있는 이름 란에 **Campaign\_Analytics\_Data** 를 입력합니다.

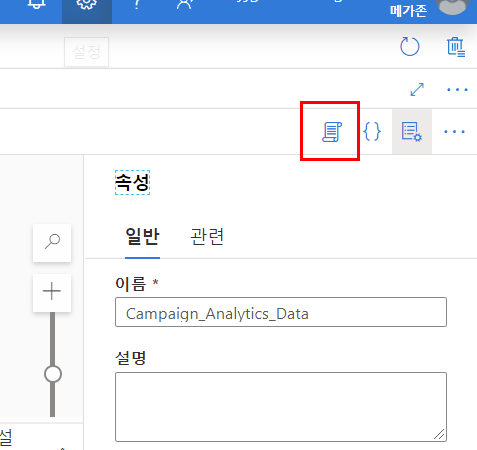


1. 화면에 보이는 “소스 추가” 를 클릭합니다.
2. 아래와 같이 Property 를 입력합니다.

|  |  |
| --- | --- |
| Field | Value |
| **출력 스트림 이름** | Enter **campaignAnalyticsCsv** |
| **원본 유형** | Select 통합 데이터 세트 |
| **데이터 세트** | Select **asamcw\_campaign\_analytics\_csv** |
| **건너뛰기 줄 수** | Enter 1 |

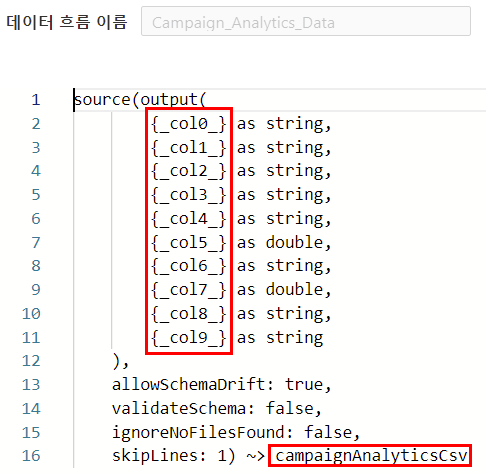


1. Data Flow 를 만들 때 데이터 미리보기 및 스키마 가져오기(프로젝션)와 같은 디버그를 설정하여 특정 기능을 사용할 수 있습니다만 시간적, 환경적 제약으로 인해 이런 과정은 생략하도록 합니다. Data Source 에는 일반적으로 정의해줘야 하는 Schema가 있으며, Data Flow의 Designer Toolbar 도구모음에서도 Script 로 구성이 가능합니다.

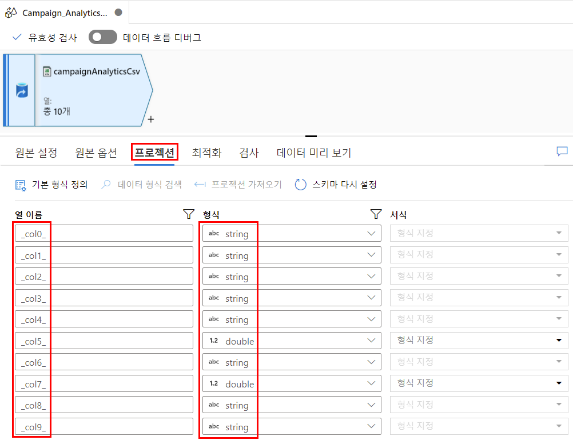


1. 스크립트 안의 내용을 아래 스크립트로 수정합니다.

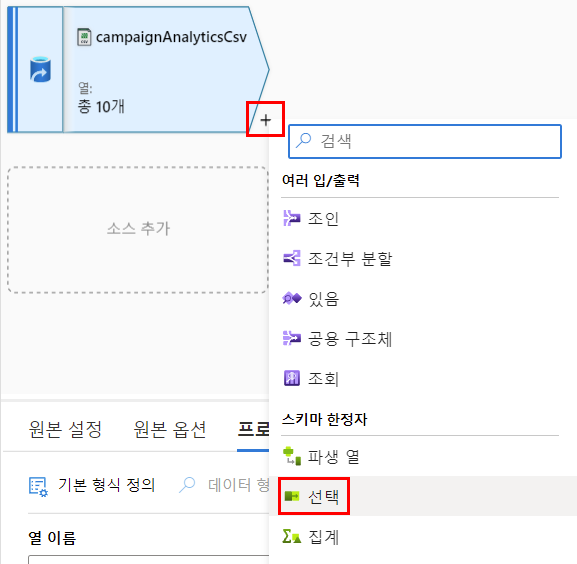
Note: 소스파일의 헤더가 잘못되어 매핑을 변경하는 시나리오입니다.



1. “프로젝션” 탭으로 이동하여 아래와 같이 나타나는지 확인합니다.

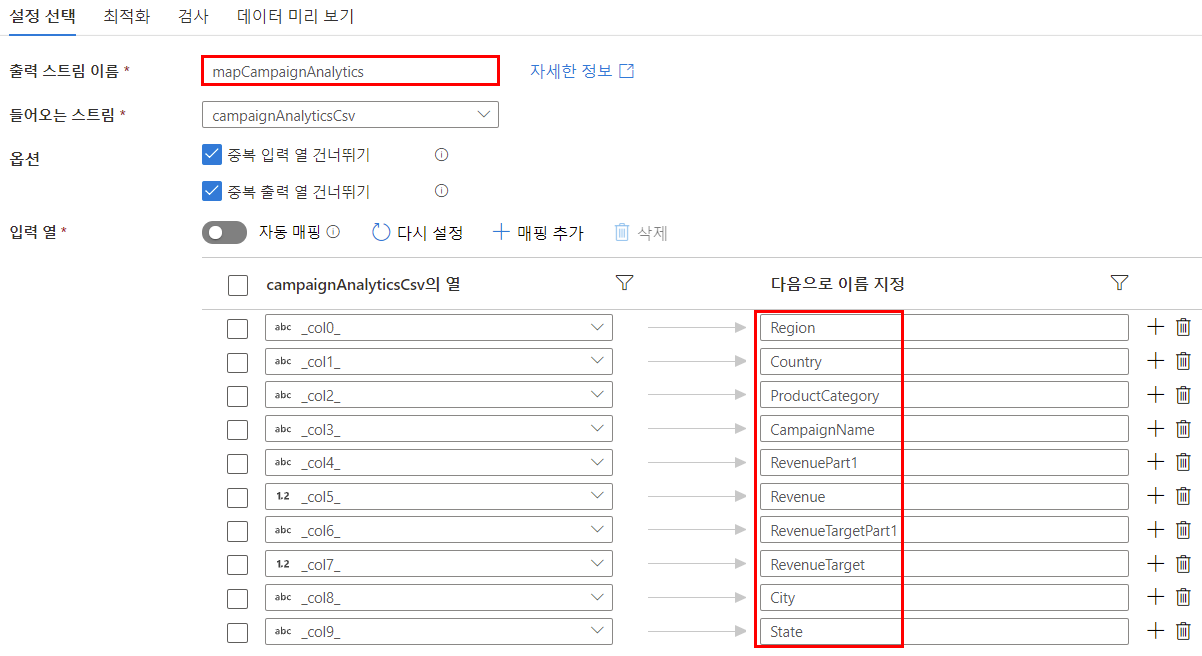


1. 이제 아래 +를 클릭하여 선택(Select)를 선택합니다.



1. 출력 스트림 이름에 **mapCampaignAnalytics** 를 입력합니다.
2. Name as 아래에 있는 입력란에 아래와 같이 순서대로 입력합니다.

* Region
* Country
* ProductCategory
* CampaignName
* RevenuePart1
* Revenue
* RevenueTargetPart1
* RevenueTarget
* City
* State

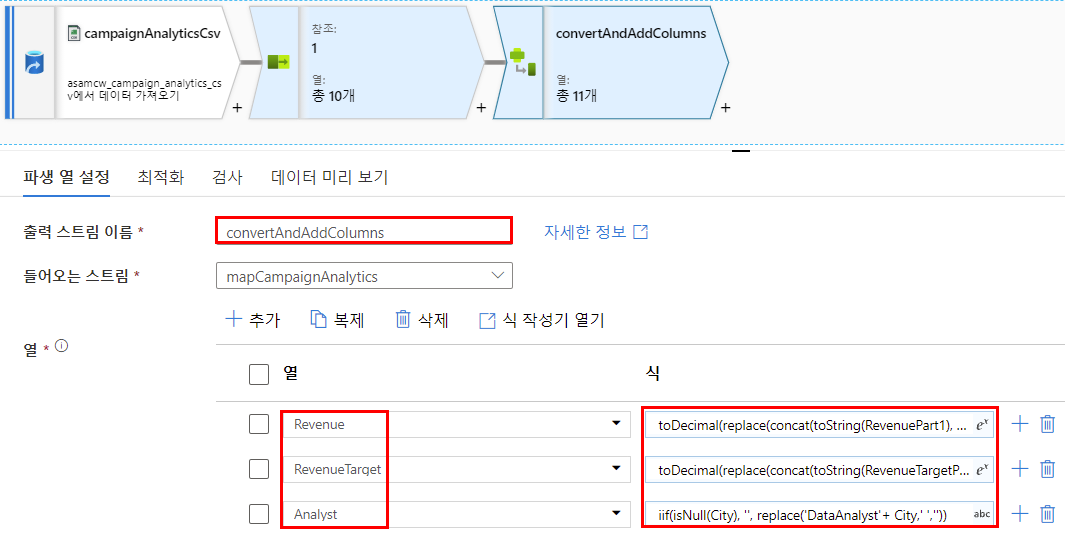


1. 다시 +를 클릭하여 파생 열(Derived Column)을 선택합니다.

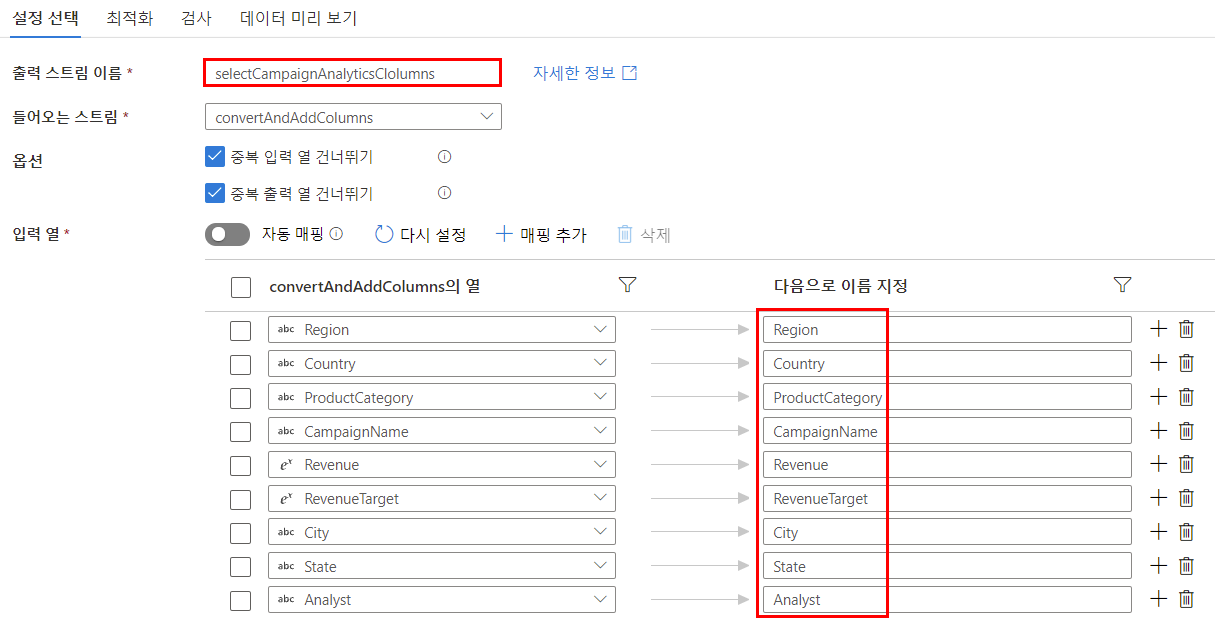


1. 이름을 **convertAndAddColumns** 로 입력합니다.
2. 아래를 참고하여 입력합니다.

|  |  |  |
| --- | --- | --- |
| Column | Expression | Description |
| **Revenue** | toDecimal(replace(concat(toString(RevenuePart1), toString(Revenue)), '\\', ''), 10, 2, '$###,###.##') | RevenuePart1과 Revenue필드를 연결하고 \ 문자가 잘못된 경우 데이터를 10진수 형식으로 변환합니다. |
| **RevenueTarget** | toDecimal(replace(concat(toString(RevenueTargetPart1), toString(RevenueTarget)), '\\', ''), 10, 2, '$###,###.##') | RevenueTargetPart1과 RevenueTarget필드를 연결하고 허용되지 않은 \ 문자를 변환한 다음, 10진수 형식으로 변환합니다. |
| **Analyst** | iif(isNull(City), '', replace('DataAnalyst'+ City,' ','')) | City필드가 NULL이면 Analyst필드에 빈 문자열을 할당하고, 그렇지 않다면 DataAnalyst를 City값에 연결하여 모든 공백구간을 제거합니다. |



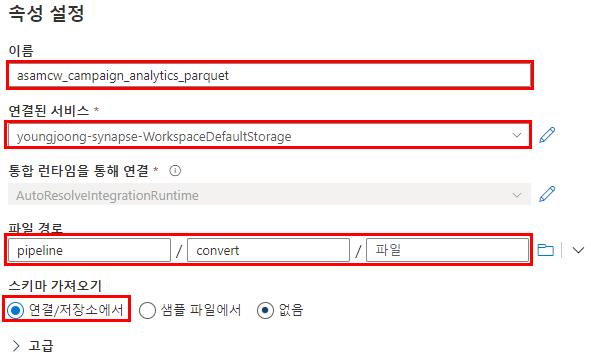
1. 다시 +를 클릭하고 선택(Select)을 선택합니다.
2. 이름에 **selectCampaignAnalyticsColumns**을 입력하고 **RevenuePart1**과 **RevenueTargetPart1**를 지웁니다.



1. 마지막으로 +를 클릭하고 싱크(Sink) 를 선택합니다.

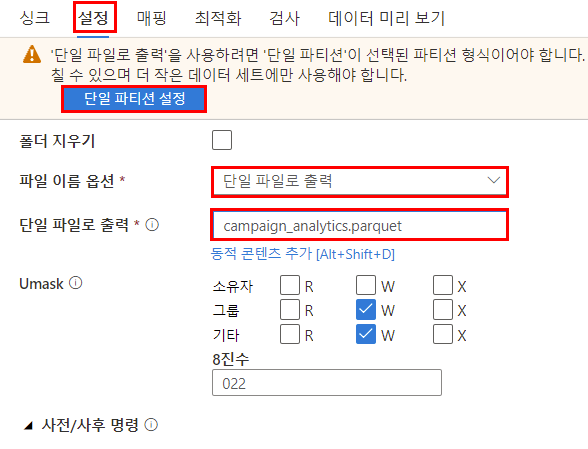


1. 이름에 **campaignAnalyticsAsa**를 입력하고 “데이터 세트”를 선택할 때   
   “+ 새로 만들기” 를 눌러 새로 생성합니다. 새 통합 데이터 세트에는 “Azure Data Lake Storage Gen2”를 선택하고, 형식 선택에는 “Parquet”를 선택합니다.
2. 이름에 **asamcw\_campaign\_analytics\_parquet** 를 입력하고 아래와 같이 속성   
   설정 값을 넣어줍니다.

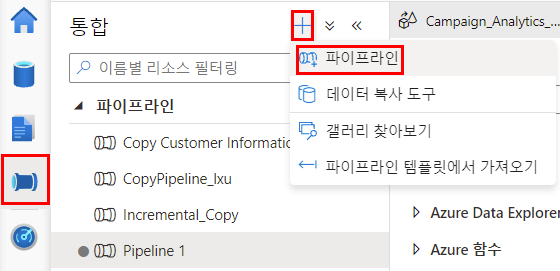


1. 위와 같이 된 것을 확인 후 “**설정**” 탭으로 가서 **파일 이름 옵션**(File Name Option)에서 **단일 파일로 출력**(Output to one single file**)**을 선택합니다.
2. 파일 명에는 **campaign\_analytics.parquet** 를 입력합니다.

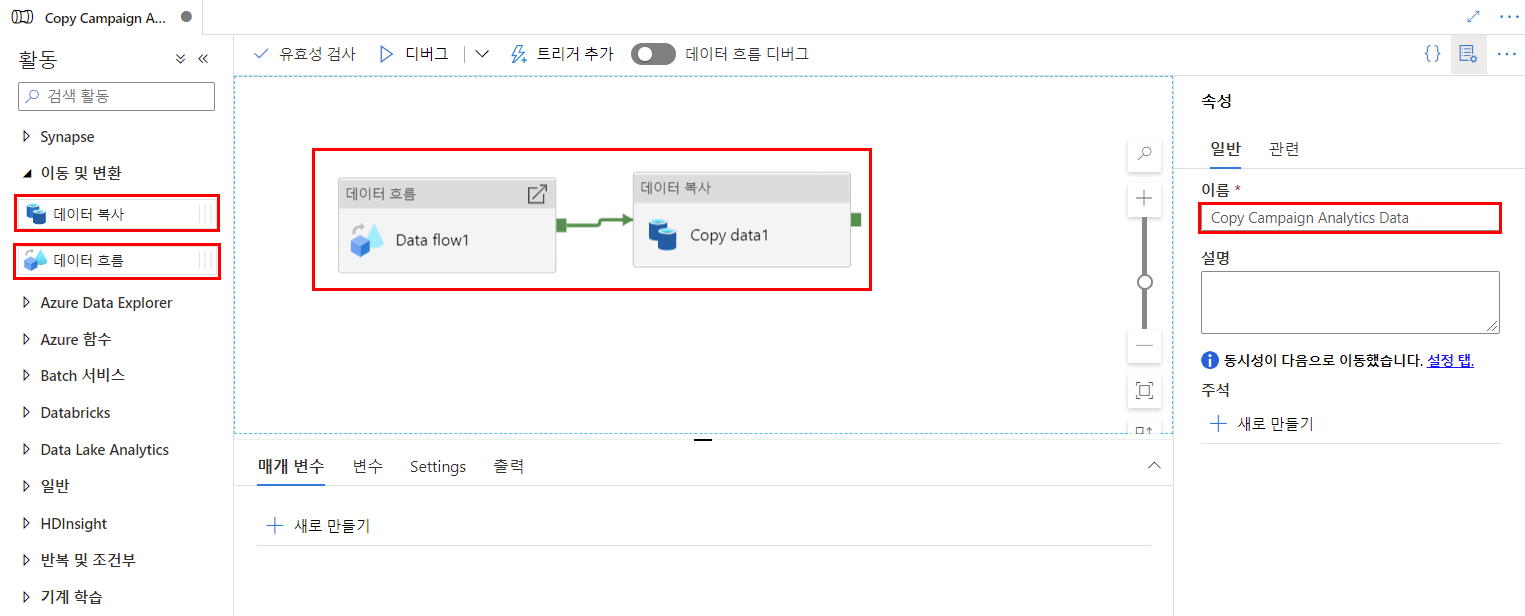
**단일 파티션 설정** (Set Single Partition)도 클릭합니다.



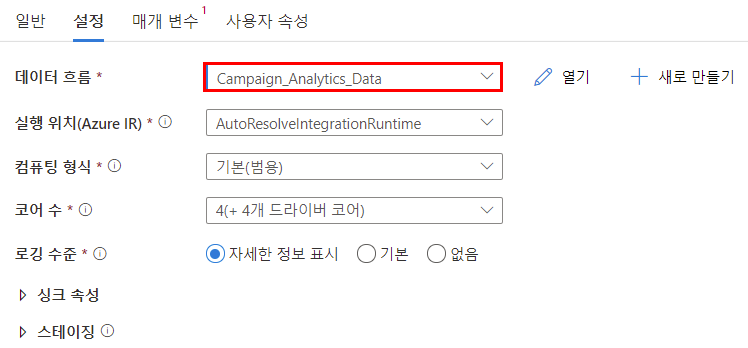
1. Publish All 하여 저장합니다.
2. 통합(Integrate) 탭에서 +를 클릭 후 파이프라인(Pipeline)을 생성합니다.



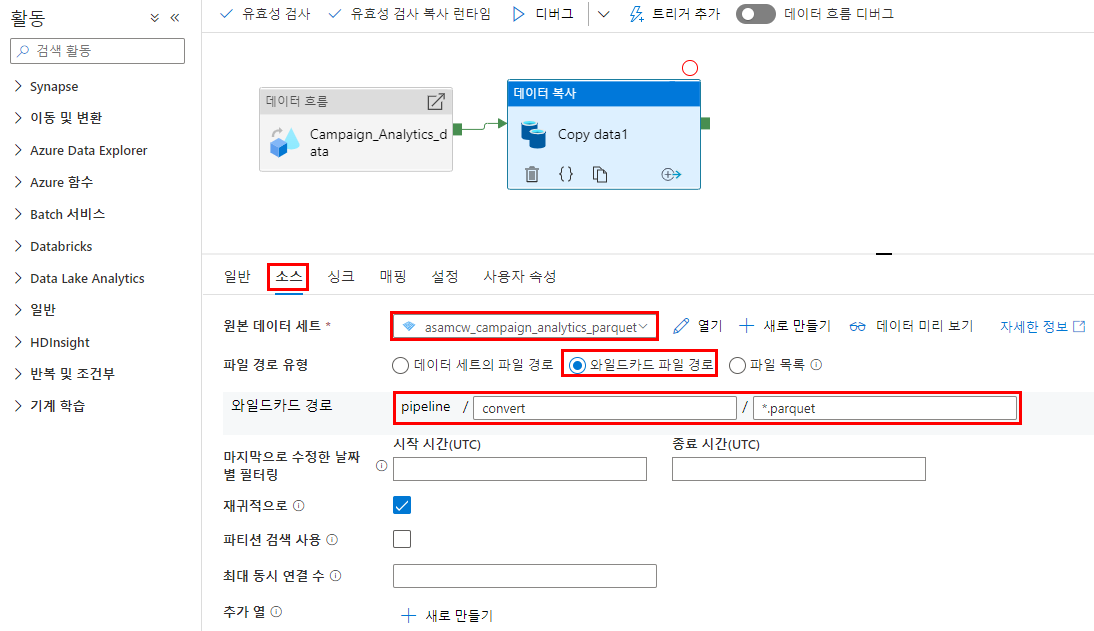
1. 속성의 이름을 **Copy Campaign Analytics Data** 로 입력합니다.
2. 이동 및 변환의 “데이터 흐름”을 Drag & Drop 하고, “데이터 복사”를 Drag & Drop 합니다. 아래와 같이 따라합니다. 데이터 흐름을 데이터 복사로 이어 줍니다.



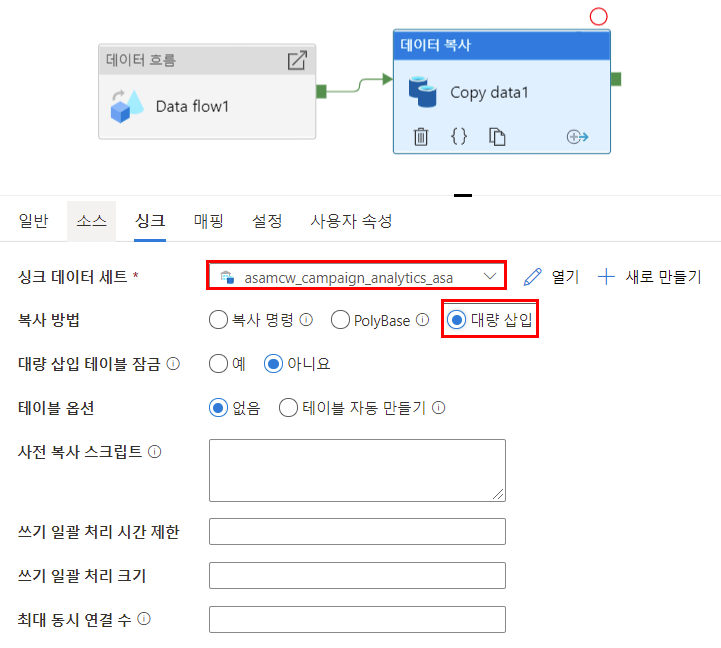
1. 데이터 흐름을 선택하고 설정 탭에 Campaign\_Analytics\_data를 선택합니다.



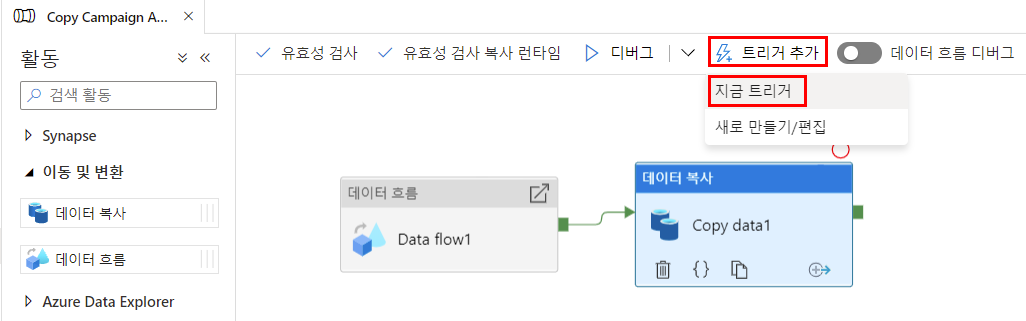
1. 데이터 복사를 선택하고, 소스(Source) 탭으로 이동하여, **asamcw\_campaignanalytics\_parquet**를 원본 소스 데이터 세트(Source Data Set) 로 선택합니다. 파일 경로 유형에 **와일드 파일 경로**(Wildcard File Path)를 선택 후 **campaign-analytics** 를 입력합니다.



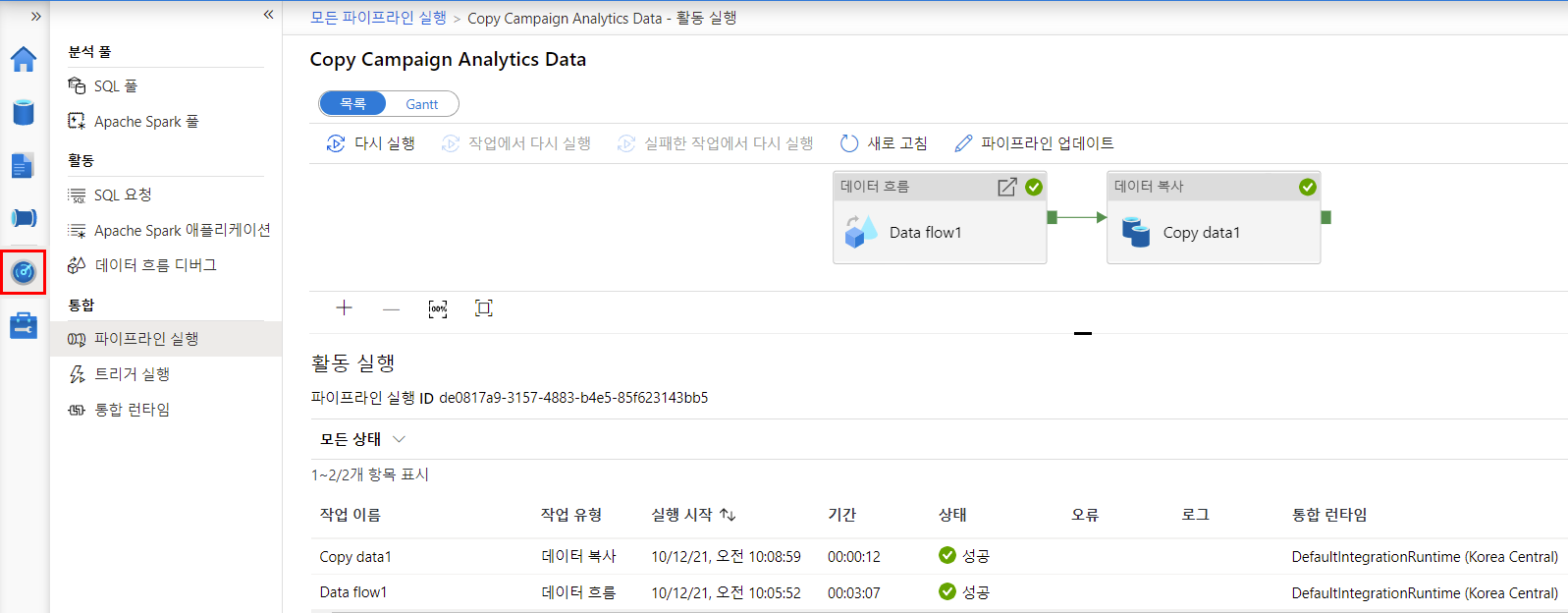
1. 싱크(Sink) 탭으로 이동해서 싱크 데이터 세트 **asamcw\_campaign\_analytics\_asa** 를 선택하고 대량 삽입(Bulk Insert)을 선택합니다.



1. 모두 게시(Publish All) 하여 저장합니다.

트리거 추가 버튼의 지금 트리거(Trigger Now)를 선택하여 실행합니다.

모니터(Monitoring) 탭에서 확인합니다.



1. 작업이 완료되면 Count\_Campaign\_Analytics\_Table 스크립트를 실행하여 결과를 확인합니다.

