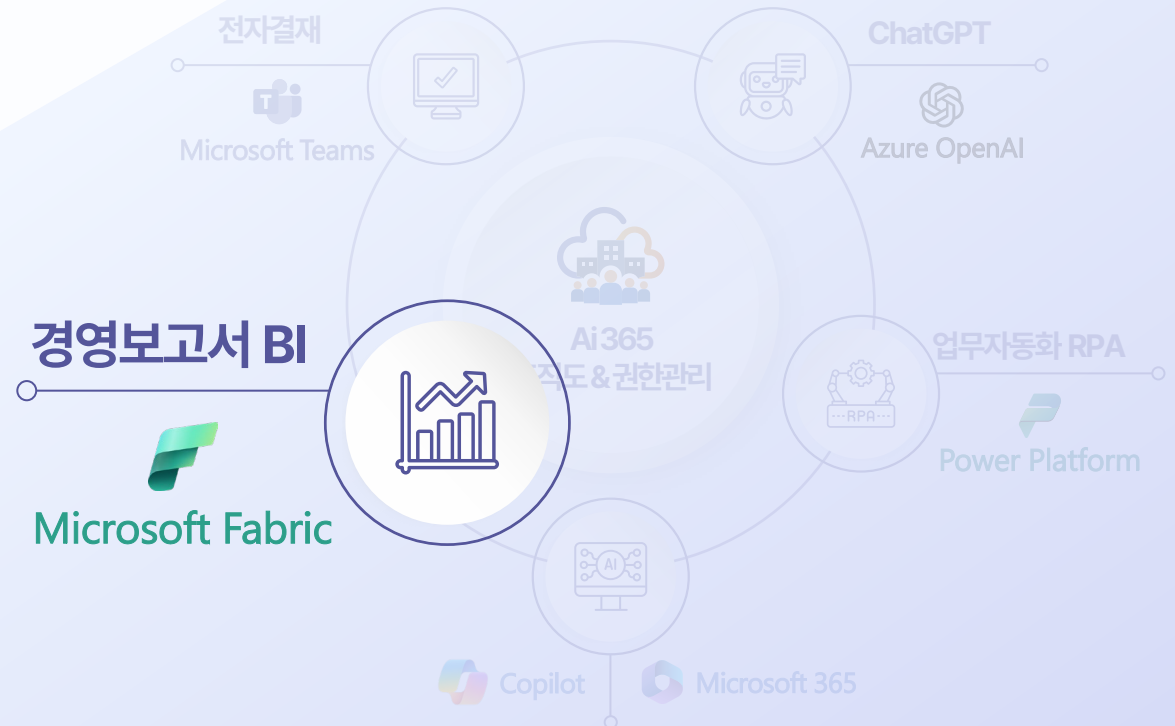


Microsoft Fabric & Power BI

Microsoft Fabric 통합 데이터 분석 플랫폼 기반 경영자 BI 소개



엠클라우드브리지
권태돈 강사

Agenda

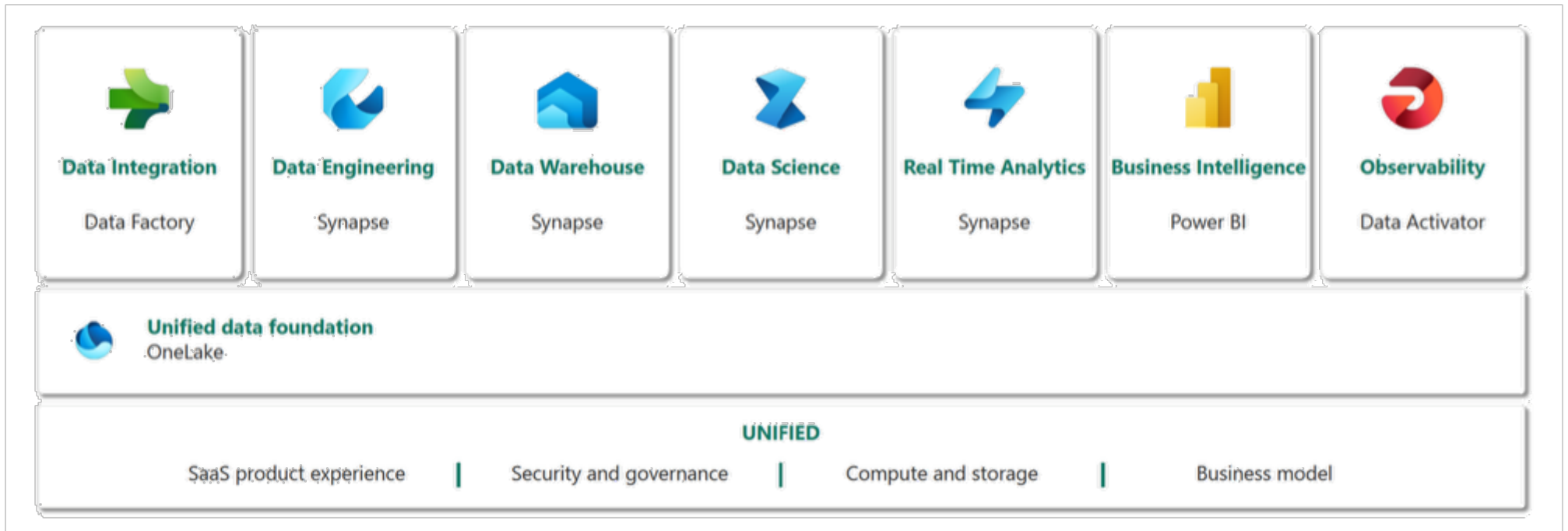
1. Microsoft Fabric 소개
2. Fabric에서 데이터 분석
3. Fabric으로 데이터 통합

1

Fabric 소개

Microsoft Fabric은 데이터 통합에서 데이터 과학, 실시간 분석 및 비즈니스 인텔리전스에 이르기까지 모든 것을 다루는 엔터프라이즈용 올인원 분석 솔루션입니다. 데이터 레이크, 데이터 엔지니어링 및 데이터 통합을 포함한 서비스 제품군을 모두 한 곳에서 제공합니다.

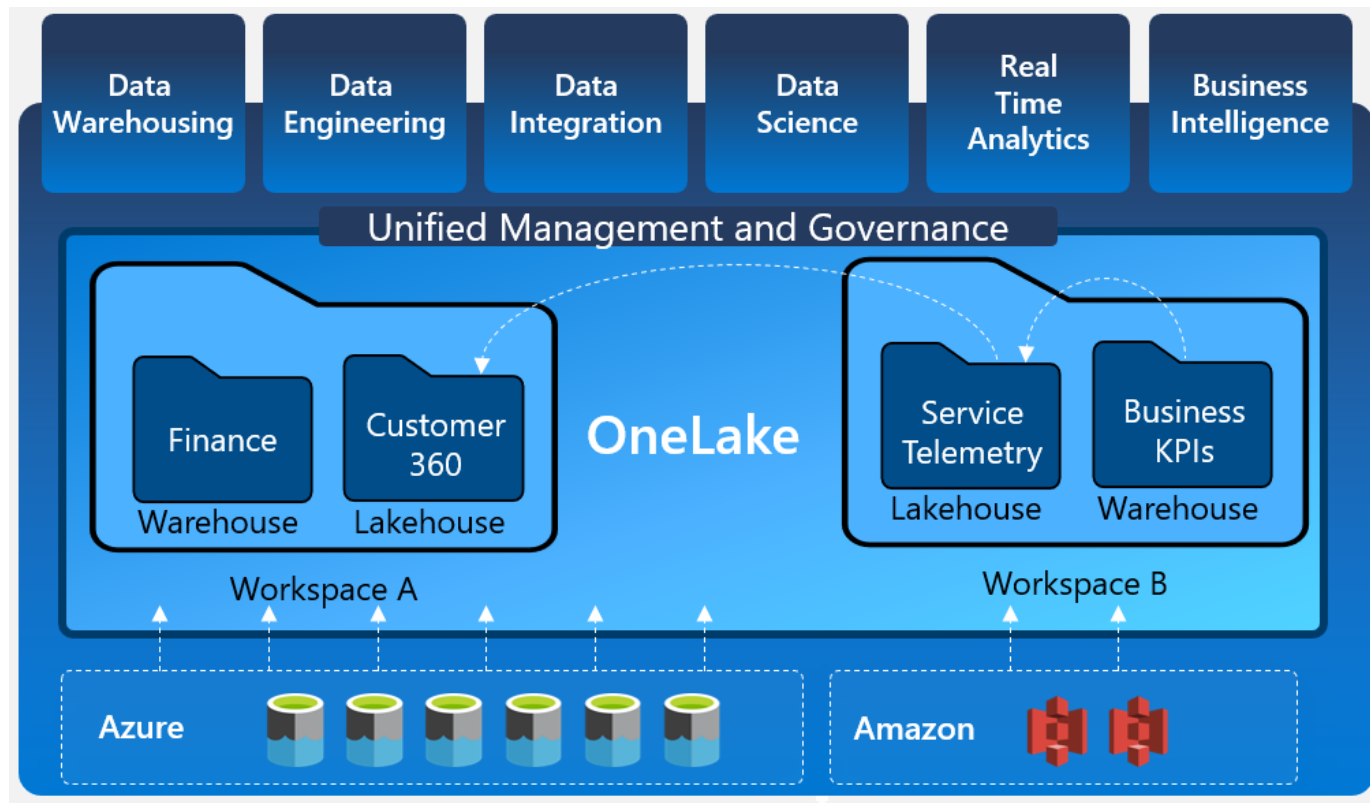
Microsoft Fabric



Source: Microsoft

Microsoft Fabric Lakehouse는 구조화되고 구조화되지 않은 데이터를 단일 위치에 저장, 관리 및 분석하기 위한 데이터 아키텍처 플랫폼입니다. 다른 데이터 관리 및 분석 도구와 통합되어 데이터 엔지니어링 및 분석을 위한 포괄적인 솔루션을 제공합니다.

OneLake - Lakehouse



Source: Microsoft

2

Fabric에서 데이터 분석

데이터 분석 환경



SQL

- SQL 분석 엔드포인트
- 웹 SQL 쿼리 편집기
- 시각적 쿼리 편집기
- SQL 편집기 연결 지원
- T-SQL
- SparkSQL
- 개체 수준 보안



Notebook

- Apache Spark
- Python, R, Scala
- SparkSQL
- 머신러닝, 딥러닝
- Synapse ML
- MLflow
- 데이터 랭글러



Power BI

- 데이터 모델
- Direct Lake
- 데이터 활성화기
- Power Query
- 페이지를 매긴 보고서
- 데이터 흐름

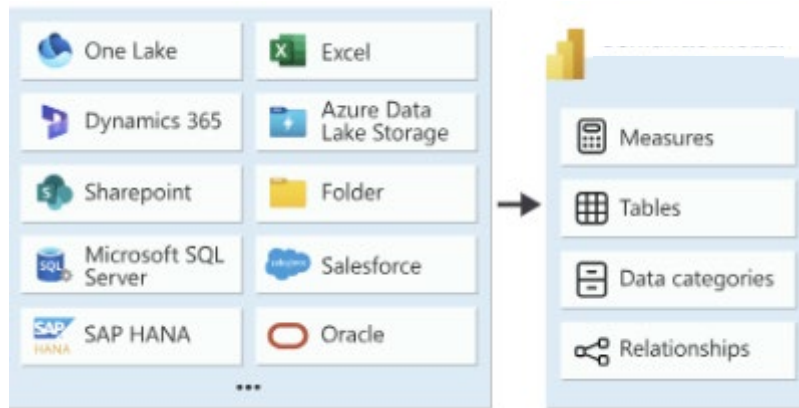


Real-time

- Eventstream
- KQL 데이터베이스
- KQL 쿼리 세트
- 실시간 대시보드

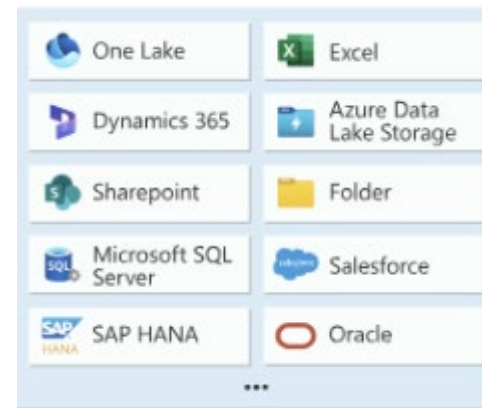
Microsoft Fabric에서 Power BI 의미 체계 모델은 분석 작업의 기본 메트릭, 비즈니스 친화적인 용어 및 표현을 사용하여 심층 분석을 가능하게 하는 논리적 설명입니다. Power BI 의미 체계 모델은 보고, 시각화, 검색 및 사용을 위해 준비된 데이터 원본을 나타냅니다.

Power BI 의미 체계 모델

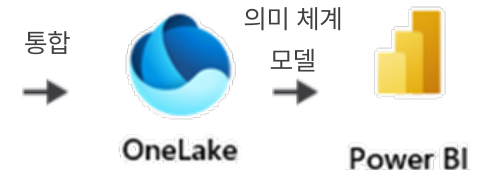


데이터 원본

Power BI



데이터 원본



Direct Lake 모드는 Power BI에서 매우 큰 데이터 볼륨을 분석하기 위한 획기적인 새 데이터 세트 기능입니다. Direct Lake는 Power BI 데이터 세트에 데이터를 가져오거나 복제할 필요 없이 데이터 레이크에서 직접 parquet 형식의 파일을 로드하는 것을 기반으로 합니다.

Power BI Direct Lake 모드

“Direct Query Mode”

Slow, but real time



“Import Mode”

Latent & duplicative but fast



“Direct Lake Mode”

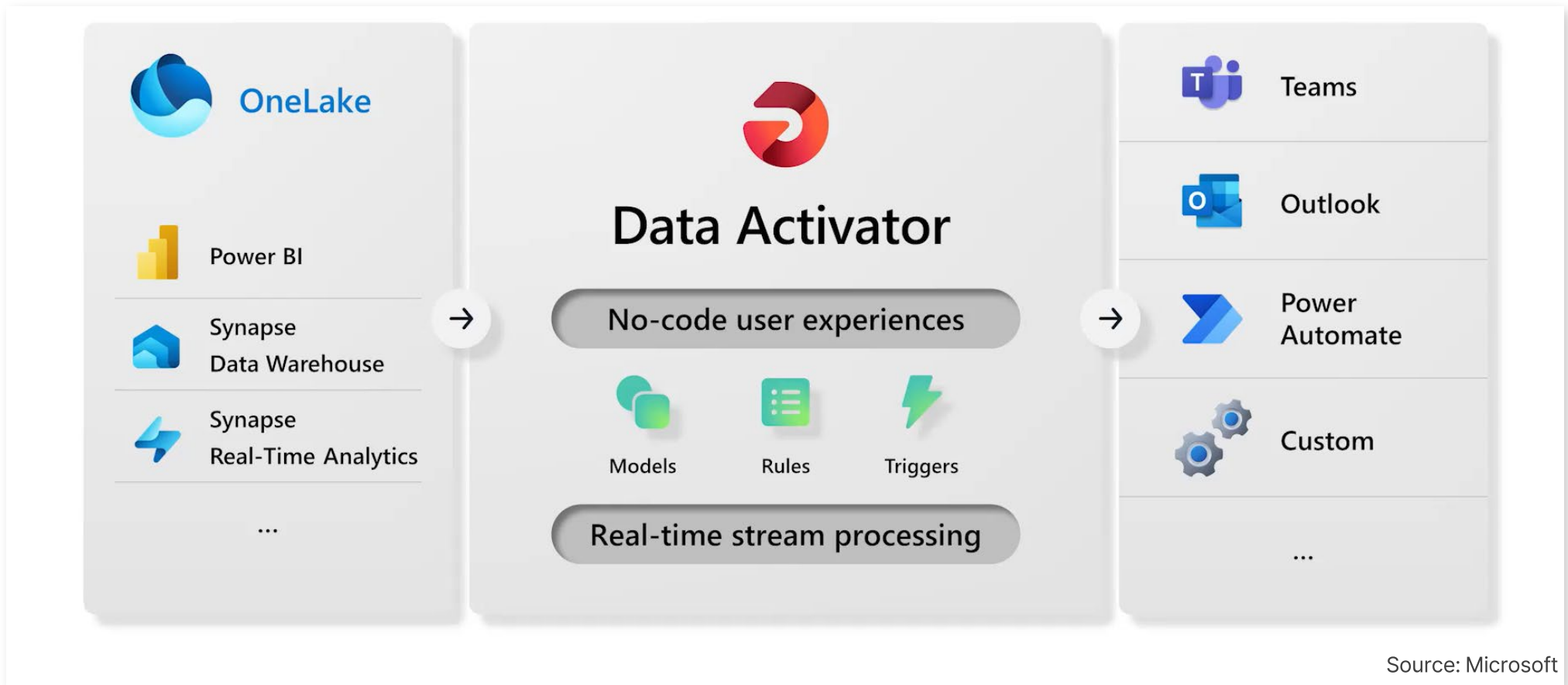
Perfect!



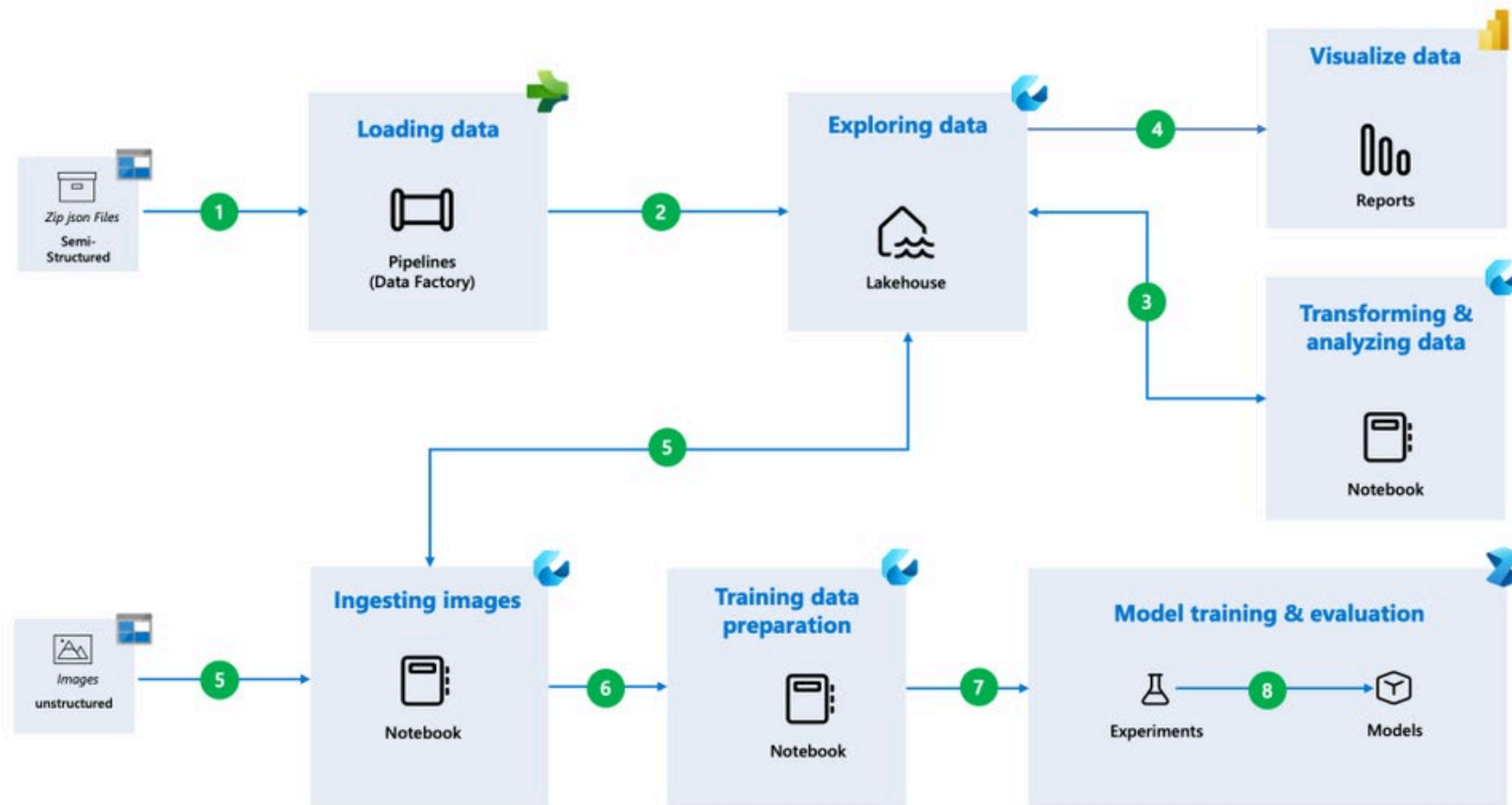
Source: Microsoft

데이터 활성화기(Data Activator)는 변경 데이터에서 패턴 또는 조건이 검색될 때 자동으로 작업을 수행하는 Microsoft Fabric의 코드 없는 환경입니다. 데이터가 특정 임계값에 도달하거나 다른 패턴과 일치하는 경우 Power BI 보고서 및 Eventstreams 항목을 모니터링합니다.

데이터 활성화기



데이터 과학을 위한 Fabric



노트북에서 데이터 액세스

```
# 파일에서 데이터 로드
df = spark.read.parquet('Files/Employee/employee.parquet')
df = pd.read_parquet('/lakehouse/default/' + 'Files/Employee/employee.parquet')

# 테이블에서 데이터 로드
df = spark.read.table('northwind.Orders')

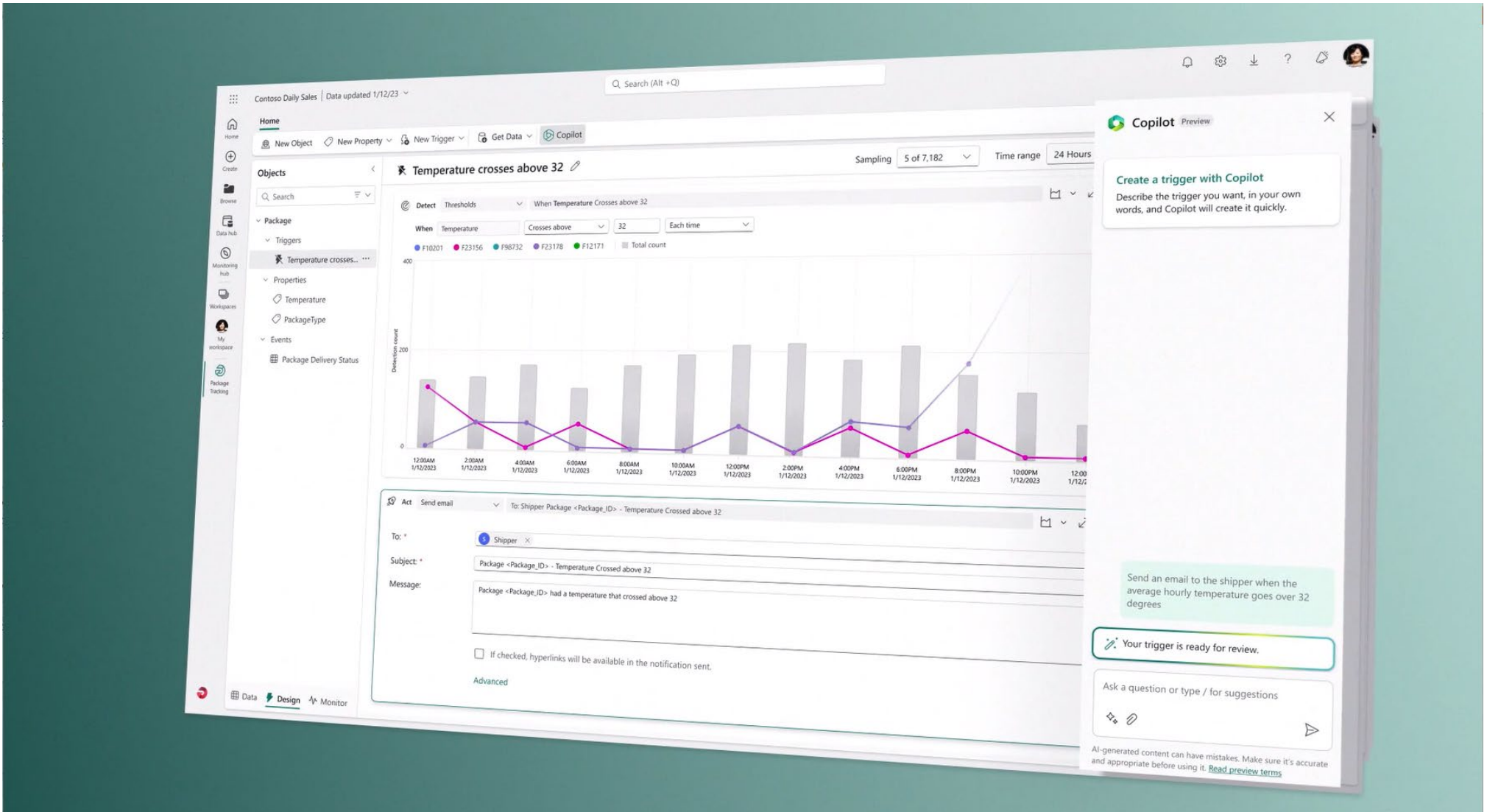
# SQL문을 이용한 데이터 로드
df = spark.sql('select * from Orders limit 10')

%%sql
select * from Orders limit 10

# Power BI 데이터세트에서 데이터 로드
import sempy.powerbi as pbi

dataset_name = 'orders'
df = pbi.read_table(dataset_name, 'State')
```

Fabric 코파일럿



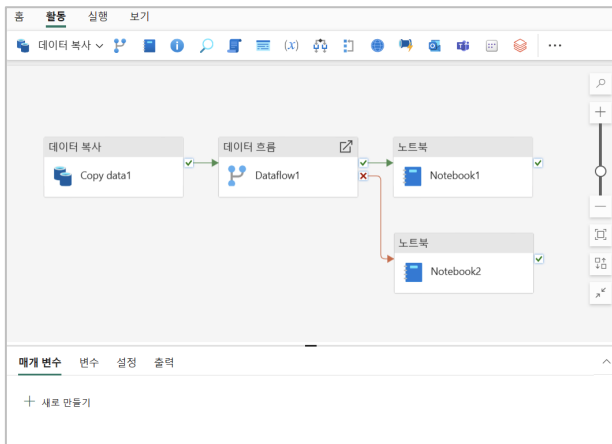
3

Fabric으로 데이터 통합

Microsoft Fabric의 데이터 엔지니어링을 통해 사용자는 조직에서 대량의 데이터를 수집, 저장, 처리 및 분석할 수 있는 인프라 및 시스템을 설계, 빌드 및 기본 수 있습니다. 데이터 통합에 사용할 수 있는 주요 도구는 데이터 파이프라인, 데이터 흐름 2세대, 노트북이 있습니다.

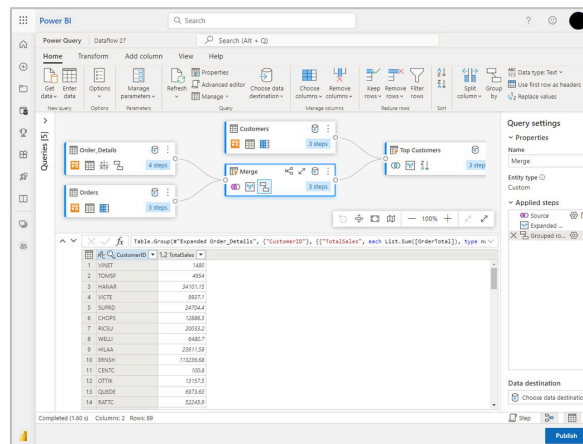
데이터 통합을 위한 주요 도구

데이터 파이프라인



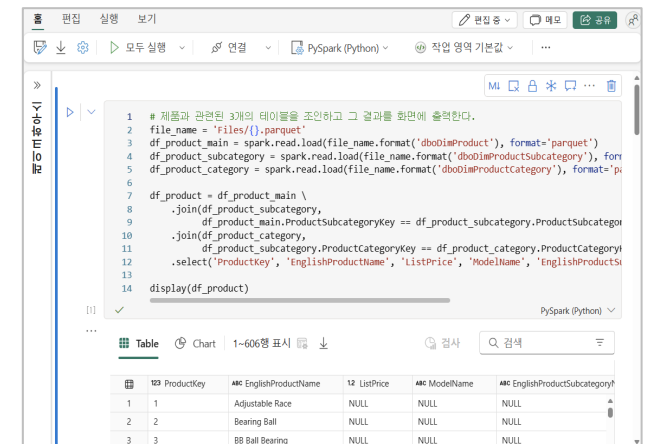
데이터 파이프라인은 클라우드에서 대규모 데이터를 처리하고, 복잡한 ETL 워크플로와 조건부 로직을 통합하여 효율적인 데이터 관리를 가능하게 합니다.

데이터 흐름 2세대



데이터 흐름은 로우 코드 인터페이스를 통해 다양한 데이터 소스에서 데이터를 수집하고, 300개 이상의 변환을 사용해 데이터를 처리한 후, 로드합니다.

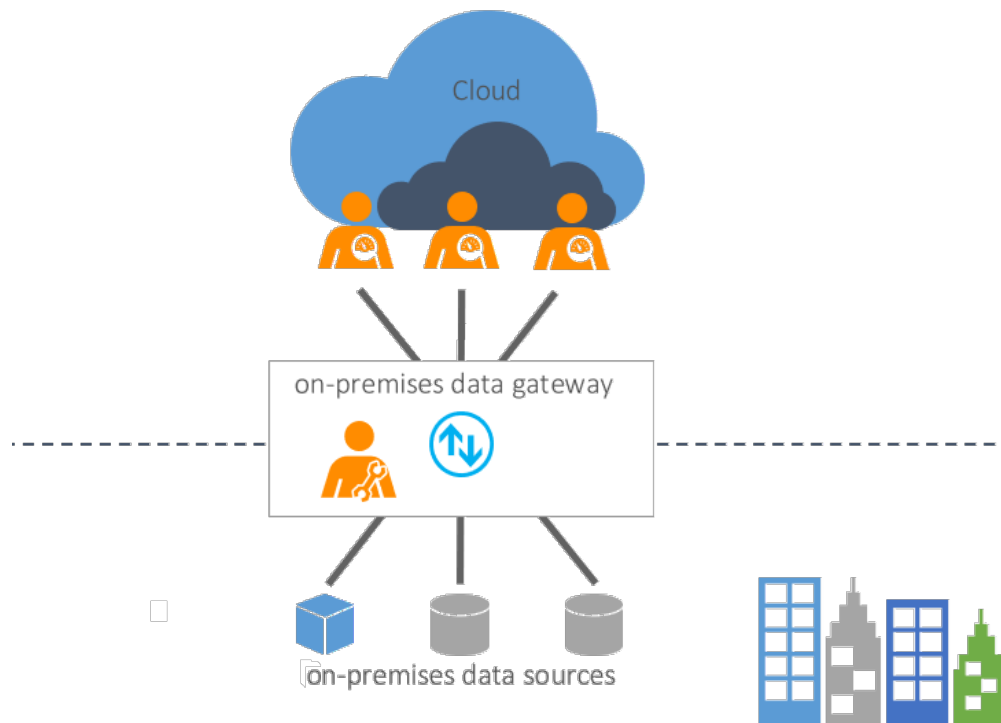
노트북



노트북은 Spark 작업을 개발하기 위한 기본 코드 항목입니다. 데이터 엔지니어는 데이터 수집, 데이터 준비 및 데이터 변환을 위한 코드를 작성합니다.

데이터 원본이 온-프레미스에 있는 시나리오에서 온-프레미스 환경과 클라우드 간의 격차를 안전하게 해소하기 위해 온-프레미스 데이터 게이트웨이를 사용할 수 있습니다. 클라우드 서비스와 온-프레미스 데이터 소스 간의 데이터 연결을 가능하게 하는 중요한 도구입니다.

온-프레미스 게이트웨이 지원



- 데이터 통합: Microsoft Fabric과 온-프레미스 데이터 소스(예: SQL Server, Oracle) 간의 안전한 데이터 연결을 제공합니다.
- 간편한 설치 및 관리: 사용자가 쉽게 설치할 수 있으며, 중앙 집중식 관리를 통해 여러 데이터 소스를 효율적으로 관리할 수 있습니다.
- 보안 데이터 전송: 데이터는 암호화되어 전송되며, 게이트웨이는 방화벽 뒤에 안전하게 배치될 수 있어 추가 보안을 제공합니다.
- 스케줄링 및 자동화: 데이터 새로고침을 예약하거나 자동화하여 최신 데이터로 작업할 수 있습니다.
- 다양한 데이터 소스 지원: 다양한 데이터베이스 및 파일 시스템을 지원하여 여러 환경에서의 데이터 접근성을 보장합니다.

Microsoft OneLake의 바로 가기(Shortcuts)를 사용하면 전체 기업을 위한 단일 가상 데이터 레이크를 생성하여 도메인, 클라우드 및 계정 전반에서 데이터를 통합할 수 있습니다. 바로 가기는 다른 스토리지 위치를 가리켜 클라우드 및 계정 전반에서 데이터를 통합할 수 있습니다.


OneLake의 바로 가기

새 바로 가기

바로 가기를 사용하여 내부 및 외부 위치에서 레이크 하우스, 웨어하우스 또는 데이터 세트에 데이터를 빠르게 끌어오세요. 항목에서 바로 가기를 업데이트하거나 제거할 수 있지만 이러한 변경 내용은 원본 데이터와 원본에 영향을 미치지 않습니다.

내부 원본

Microsoft OneLake



패브릭


외부 원본

Amazon S3




AWS

Amazon S3 호환(미리 보기)



Generic connector

Azure Data Lake Storage Gen2



Azure

× 기존 연결 새 연결 만들기

연결 설정

URL

예: https://s3.amazonaws.com

연결 자격 증명

연결

새 연결 만들기

연결 이름

https://s3.amazonaws.com

인증 종류

기본

사용자 이름

암호

 OpenAI |  Ai 365 |  Copilot

Microsoft AI 플랫폼 활용 AI 업무 환경 구축 세미나

Thank You

데이터에 가치를 더하여 고객의 성장에 공헌합니다.
Specialized Consulting Firm in **Data & AI** Cloud System