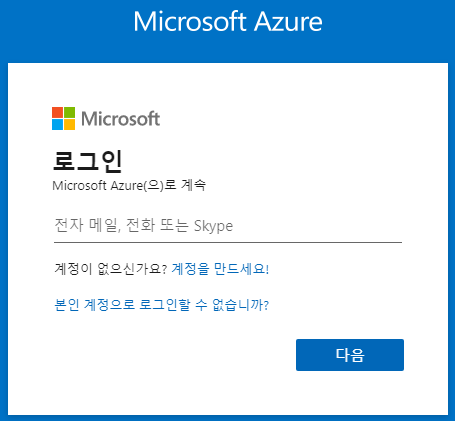
**Lab 0 – 실습 환경 설정**

# 요구 사항

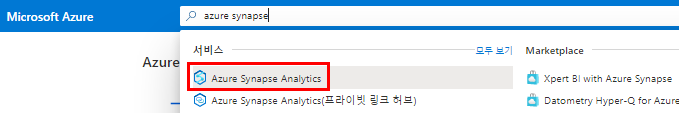
1. Microsoft Azure 구독 (Microsoft 이외의 구독은 유료 구독이어야 합니다.)

# Task 1 : Azure Synapse Workspace 생성

1. **Azure Portal**(<https://portal.azure.com/>)**에 접속**하여 개인 혹은 제공된 아이디로 **로그인** 합니다.



1. 상단의 돋보기에서 **Azure Synapse Analytics를 입력**하여 선택합니다.



1. Azure Synapse Analytics **만들기를 클릭**하여 생성할 리소스의 정보를 입력합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

1. **기본 사항**에 **필요한 정보를 입력**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

리소스 그룹 : mzc-rg (새로 만들기)

작업 영역 이름 : <username>-synapse (고유 이름값)

지역 : 한국 중부

Data Lake Storage Gen2 계정 이름 : <username>adls (새로 만들기)(고유 이름값)

Data Lake Storage Gen2 파일 시스템 이름 : synapse (새로 만들기)

1. **보안**에서 **SQL admin 관리자 계정 정보를 입력**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

SQL Server 관리자 로그인 : sqladminuser

SQL 암호 : 각자 입력

1. **검토 + 만들기**를 클릭 후 하단의 **만들기**로생성합니다.

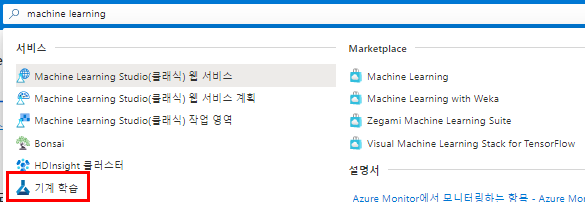
텍스트이(가) 표시된 사진

자동 생성된 설명

\* 빨간 박스 외에 리소스 설정은 기본 값이므로 수정하지 마세요.

# Task 2 : Machine Learning 생성

1. 상단의 돋보기에서 **Machine Learning을** **입력**하여 선택합니다.



1. 기계 학습 **만들기를 클릭**하여 생성할 리소스의 정보를 입력합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

1. **기본**에서 **필요한 정보를 입력**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

리소스 그룹 : mzc-rg

작업 영역 이름 : mzcmlworkspace

지역 : 한국 중부

스토리지 계정 : 작업 영역 이름 입력 시 자동 신규 생성

키 자격 증명 모음 : 작업 영역 이름 입력 시 자동 신규 생성

Application Insights : 작업 영역 이름 입력 시 자동 신규 생성

컨테이너 레지스트리 : 없음

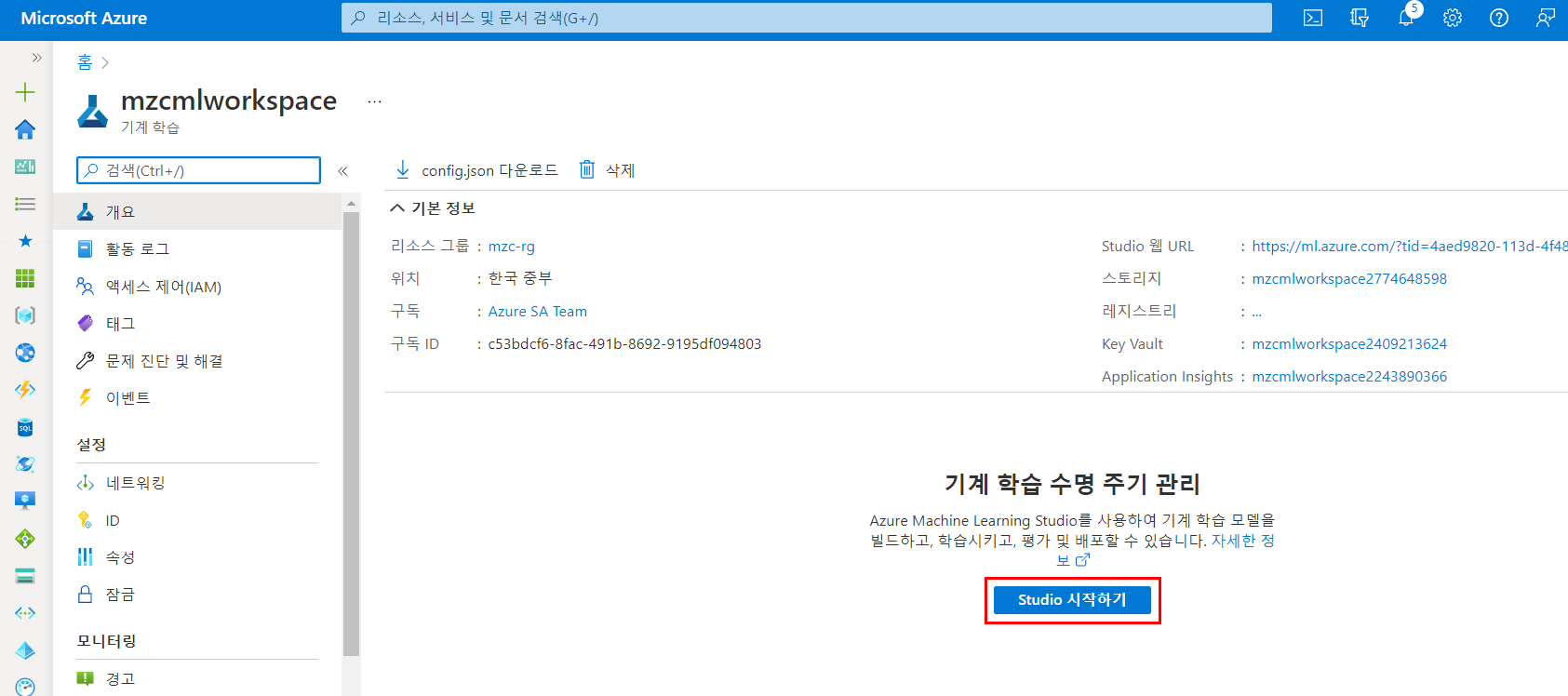
1. **검토 + 만들기**를 클릭 후 하단의 **만들기**로 생성합니다.

텍스트이(가) 표시된 사진

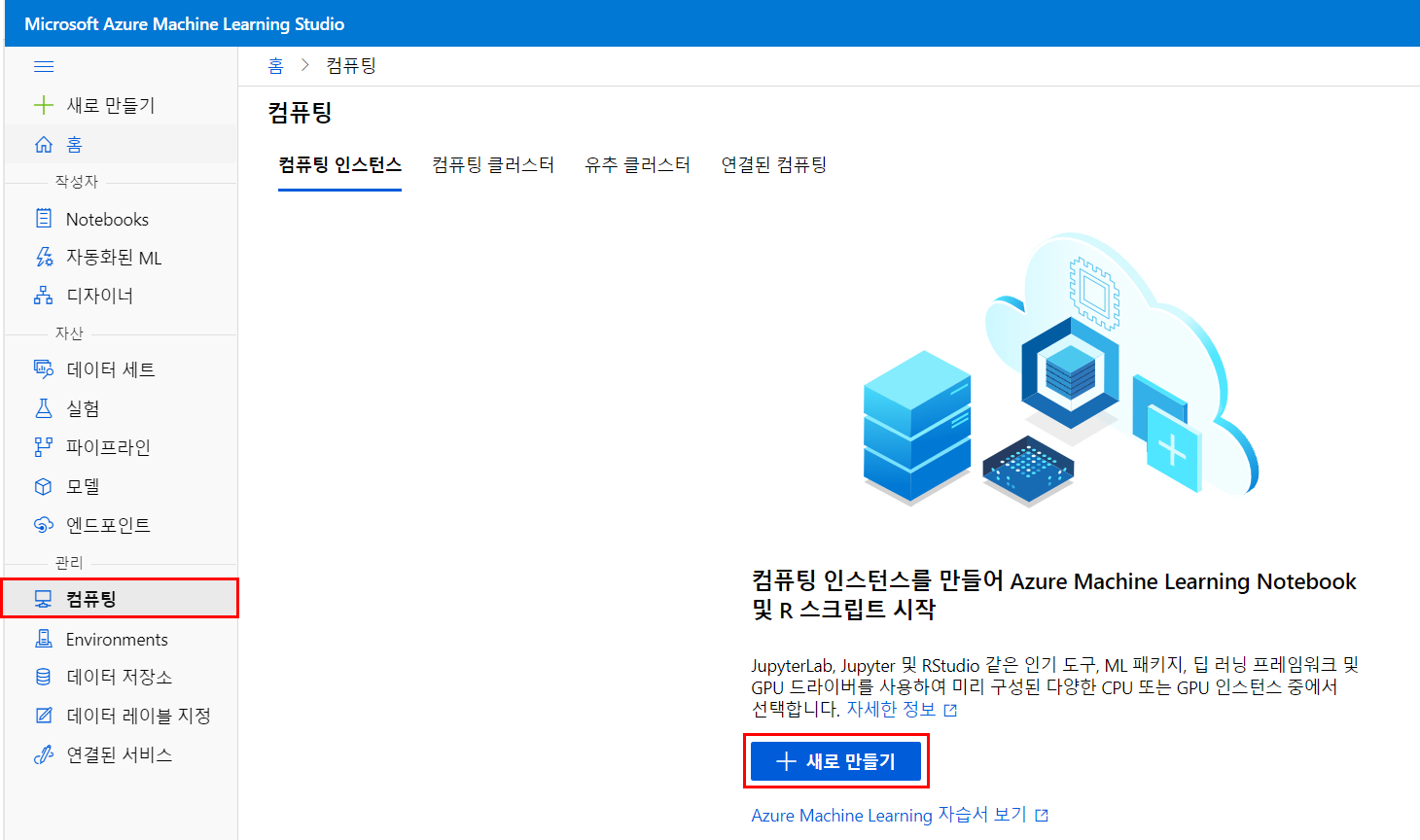
자동 생성된 설명

\* 빨간 박스 외에 리소스 설정은 기본 값이므로 수정하지 마세요.

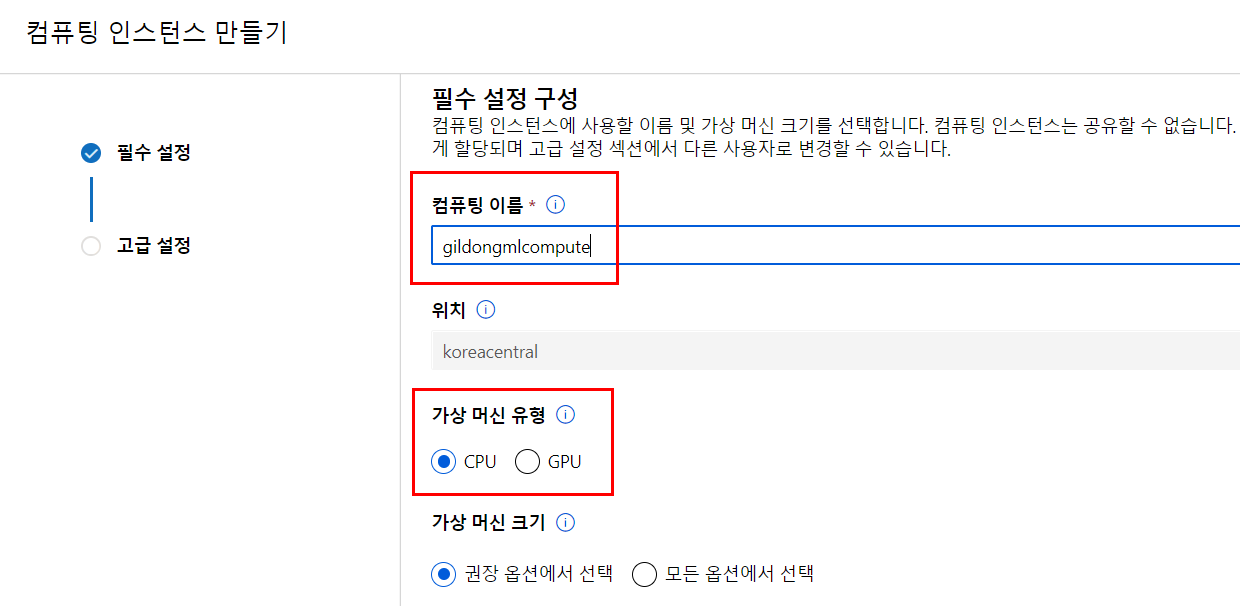
1. 방금 생성했던 **ML(Machine Learning)** 워크스페이스로 검색을 통해 이동하여 **Studio 시작하기**를 클릭합니다.



1. 접속한 스튜디오에서 왼쪽 블레이드의 **컴퓨팅**으로 이동합니다. 아래와 같이 페이지가 뜨면 **새로 만들기**를 클릭합니다.



1. **컴퓨팅 이름, 가상머신 유형, 가상 머신 크기**를 선택합니다. **만들기**를 클릭합니다.



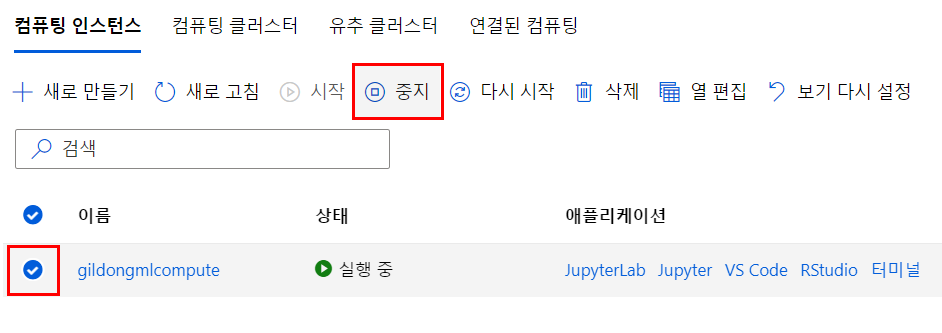


컴퓨팅 이름: <username>mlcompute

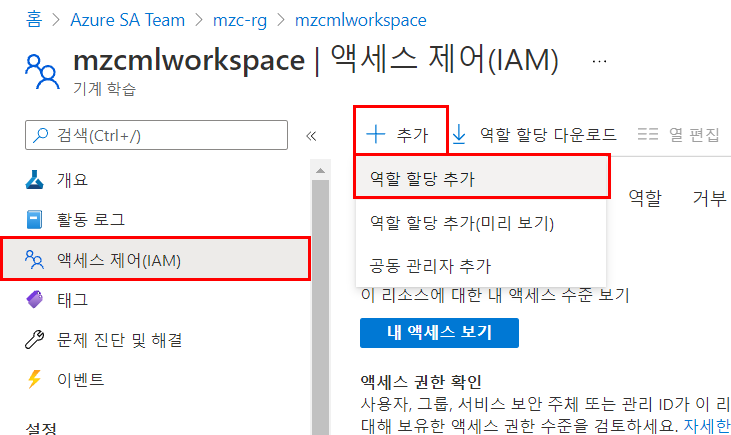
가상 머신 유형: CPU

가상 머신 크기: Standard\_DS3\_v2

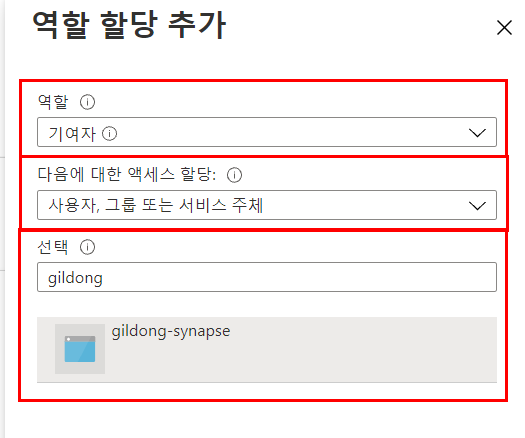
1. 컴퓨팅 인스턴스가 실행 중으로 표시됩니다. **과금 방지**를 위해 **인스턴스를 선택**하고 상단의 **중지 버튼**을 눌러 인스턴스를 중지시켜 놓습니다.



1. Studio를 닫고 **ML(Machine Learning)** 워크스페이스로 이동합니다. 왼쪽 블레이드의 **액세스 제어(IAM)**으로 이동하여 **역할 할당을 추가**합니다.

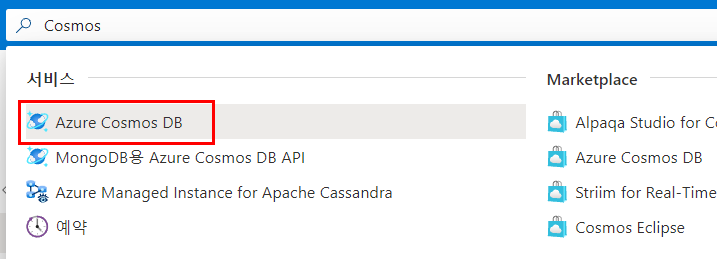


1. **역할**은 기여자, **선택** 부분에 Synapse Workspace의 이름을 검색합니다. Synapse Analytics를 선택합니다. **저장**을 눌러 역할을 할당합니다.

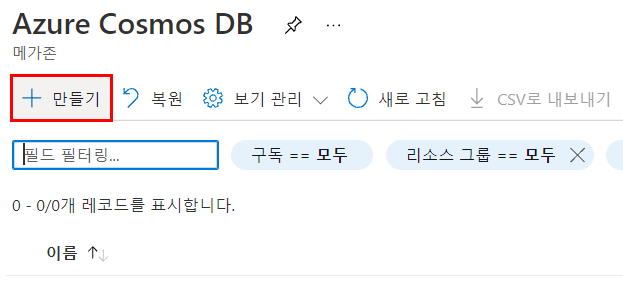


# Task 3 : Cosmos DB생성

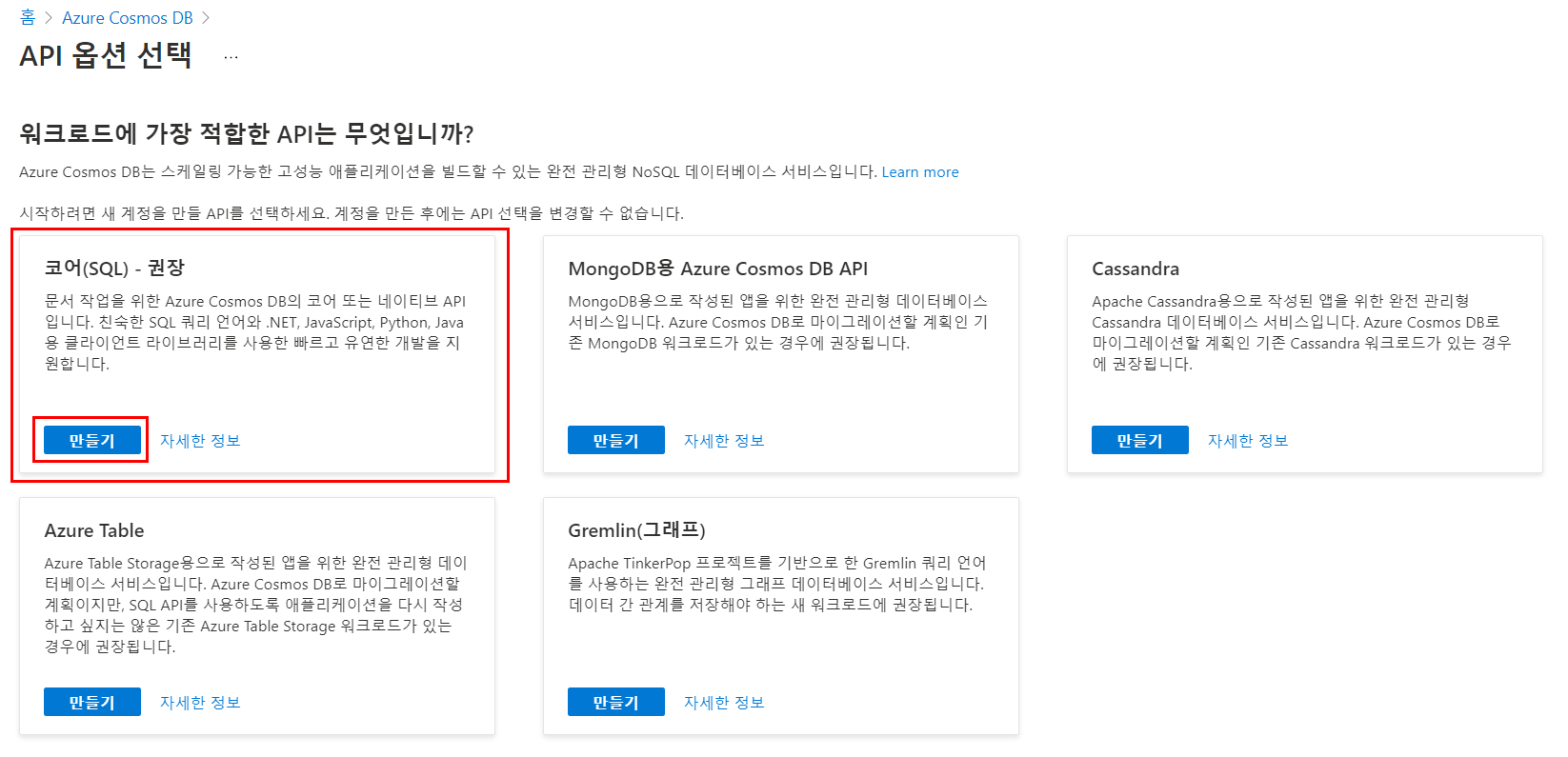
1. 상단의 돋보기에서 **Comsmos DB를** **입력**하여 선택합니다.



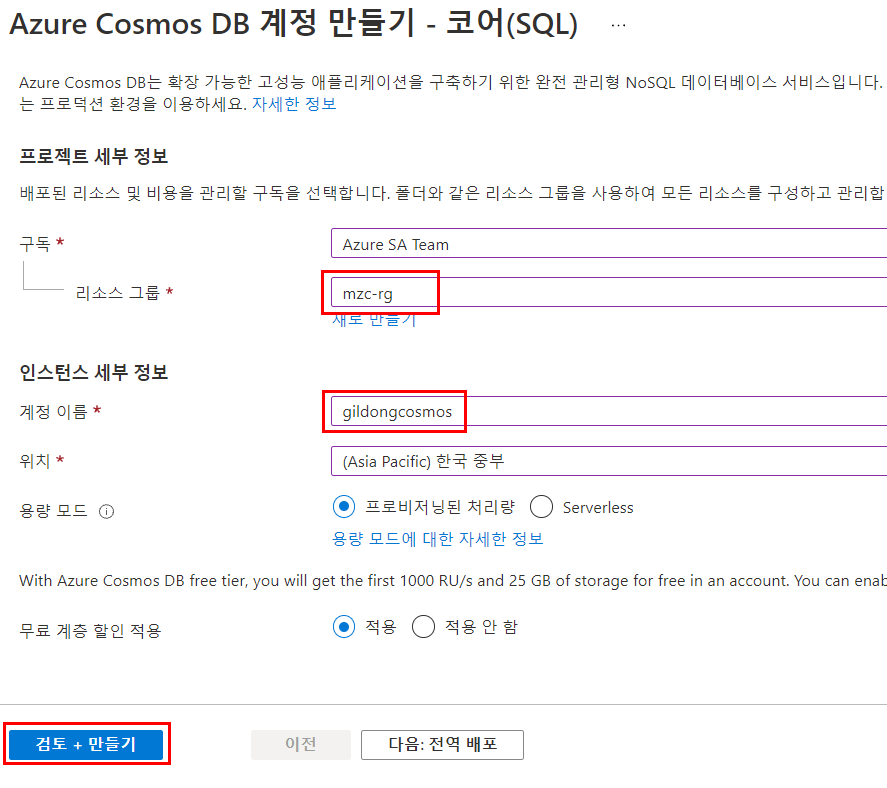
1. **Azure Cosmos DB 만들기를 클릭**하여 생성할 리소스의 정보를 입력합니다.



1. API 중 **코어(SQL) API**를 선택합니다.



1. **기본**에서 **필요한 정보를 입력**합니다. **검토 + 만들기를 클릭 후 하단의 만들기로 생성합니다.**



리소스 그룹 : mzc-rg

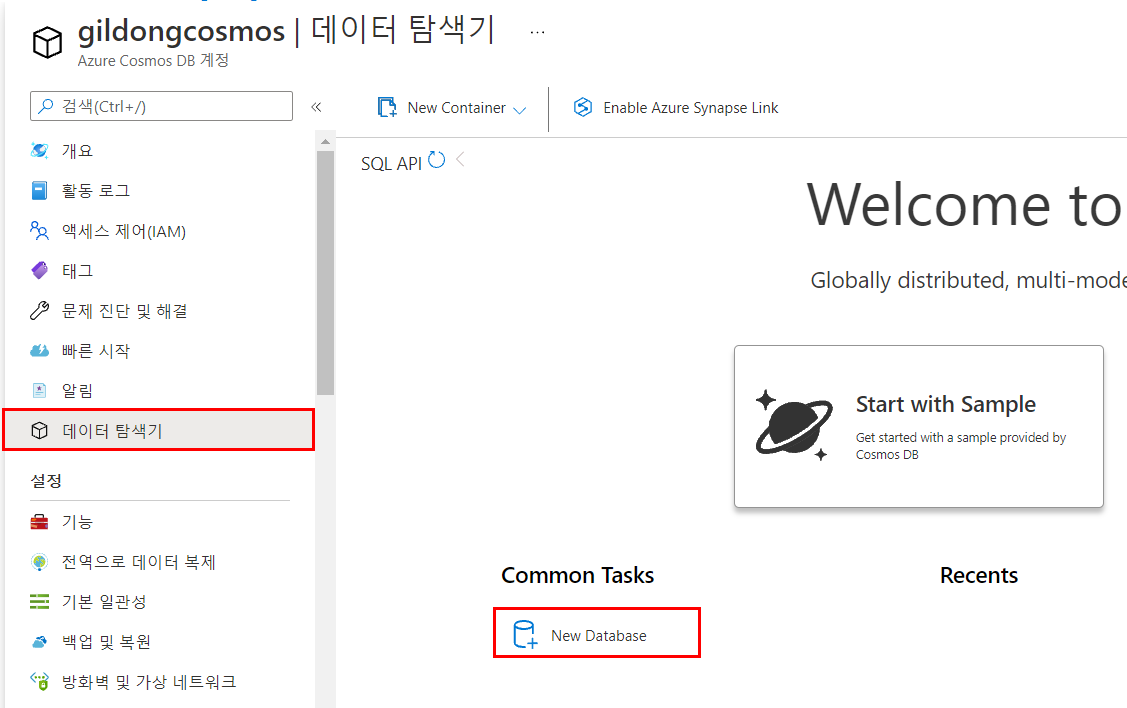
리소스 이름 : <username>cosmos (고유 이름값)

지역 : 한국 중부

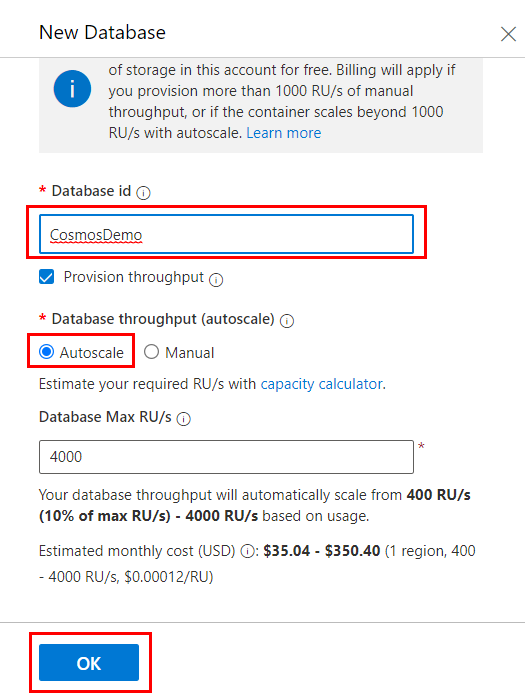
용량 모드 : 프로비저닝된 처리량

무료 계층 할인 적용 : 적용

1. 만들어진 **Cosmos DB 계정**으로 이동합니다. 왼쪽 메뉴에서 **데이터 탐색기**를 선택합니다. **New Database**를 클릭합니다.



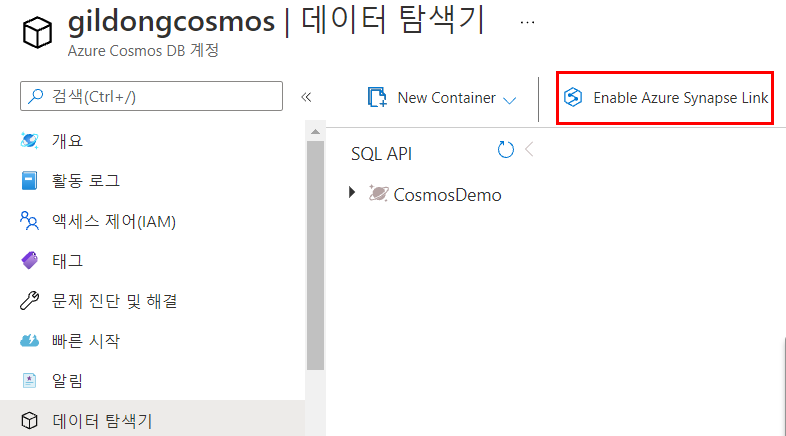
1. **Database id**를 아래와 같이 입력합니다. **Autoscale**을 설정하고 **Database Max RU/s**는 기본값인 4000으로 설정합니다. **OK**를 눌러 데이터베이스를 생성합니다.

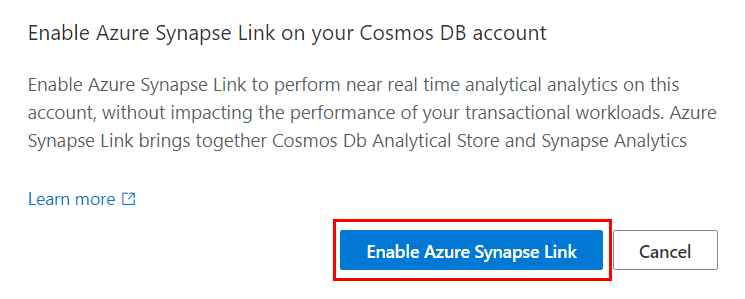


Database id: CosmosDemo

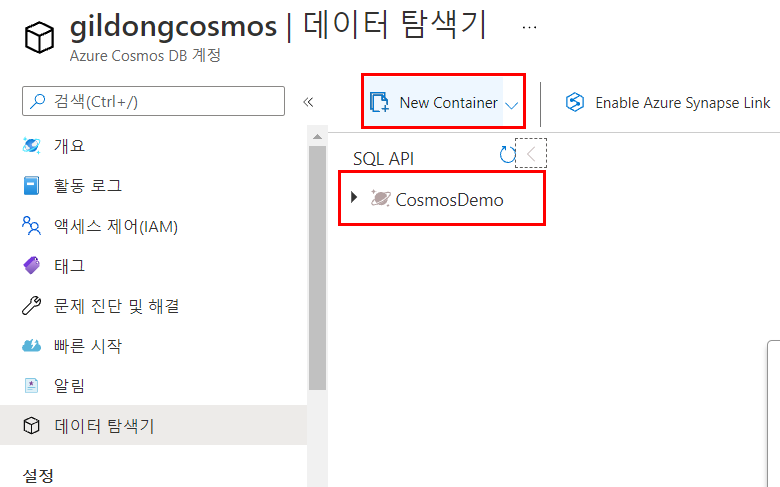
Database Max RU/s : 4000

1. 아래와 같이 **데이터베이스가 생성**된 것을 볼 수 있습니다. 상단의 **Enable Azure Synapse Link**를 클릭하여 Synapse Link를 허용합니다. (시간이 조금 소요됩니다)

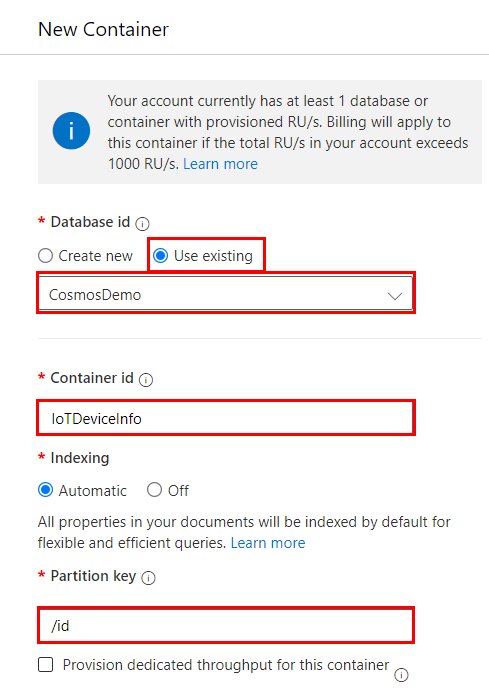




1. 상단의 **New Container**를 클릭하여 컨테이너를 생성합니다.



1. **Database id**는 **Use existing**을 선택하여 기존에 생성한 **CosmosDemo**를 선택합니다. **Container id**와 **Partition key**를 아래와 같이 생성합니다.

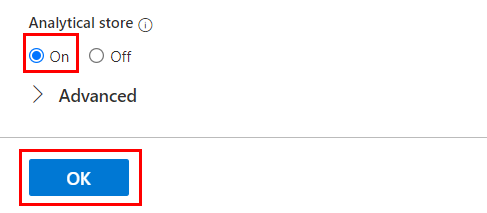


Database id : CosmosDemo

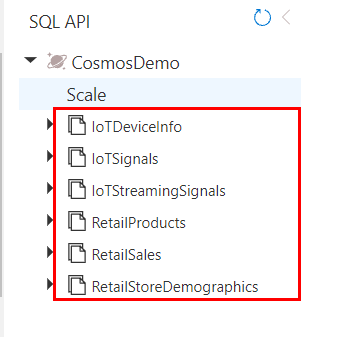
Container id : IoTDeviceInfo

Partition key : /id

1. **Analytical store**를 **On**으로 설정합니다. **OK**를 클릭하여 컨테이너를 선택합니다.

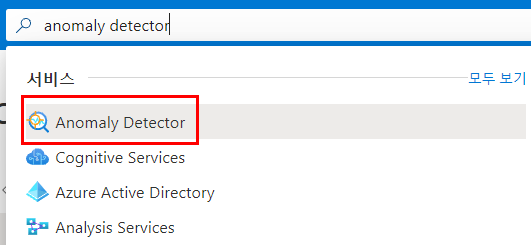


1. 같은 방법으로 **IoTSignals, IoTStreamingSignals, RetailProducts, RetailSales, RetailStoreDemographics** 컨테이너를 생성합니다. **Partition Key는 동일하게 /id를 입력**합니다. **Analytical Store**의 경우, **IoTStreamingSignals만 Off**를 선택합니다. **나머지는 On**을 선택합니다. 최종 모습은 아래와 같습니다.

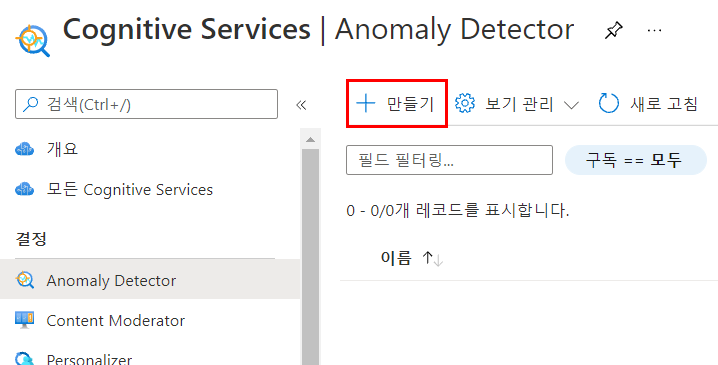


# Task 4 : Anomaly Detector (Cognitive Service) 생성

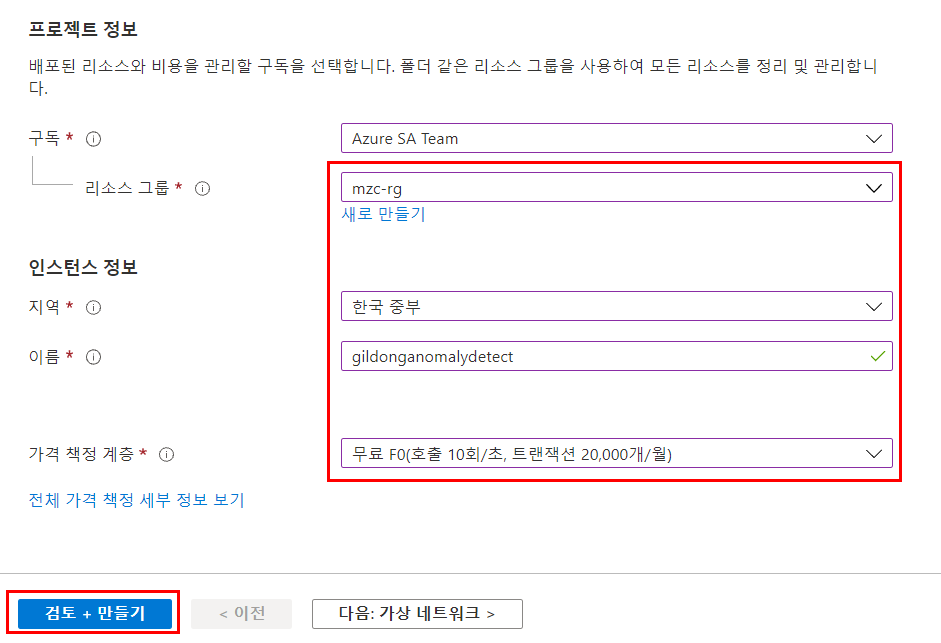
1. 상단의 돋보기에서 **Anomaly Detector를 입력**합니다.



1. **+만들기**를 클릭하여 생성합니다.



1. **기본**에서 필요한 정보를 입력합니다. **검토 + 만들기**를 클릭 후 하단의 **만들기**로 생성합니다.



리소스 그룹 : mzc-rg

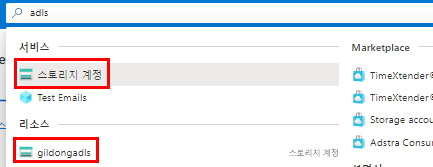
지역 : 한국 중부

리소스 이름: <username>anomalydetect

가격 책정 계층 : 무료 F0

# Task 5 : ADLS Gen2에 컨테이너 및 데이터 생성

1. 상단의 돋보기에서 **ADLS를 입력**하여 앞에서 생성한 <username>adls를 선택합니다.

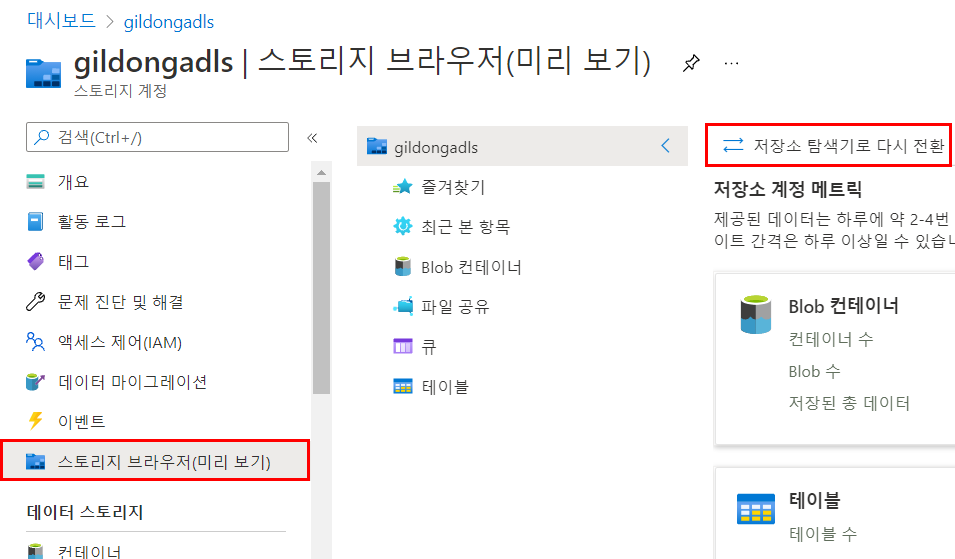


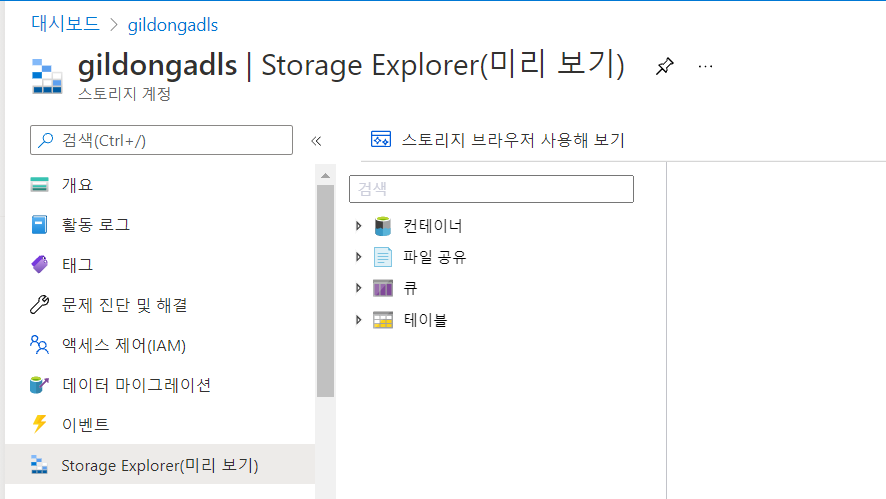
1. **왼쪽 블레이드**에서 **Storage Explorer를 클릭**하고, **컨테이너 클릭합니다.** 앞에서 생성했더니 mzcsynapse 파일 시스템이 보이는지 확인합니다.

텍스트이(가) 표시된 사진

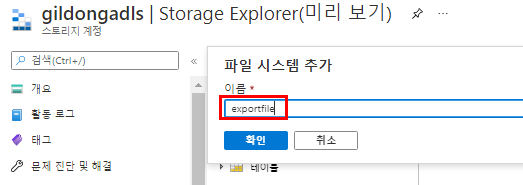
자동 생성된 설명

\*아래와 같이 **스토리지 브라우저(미리 보기)**가 적용되는 경우, 스토리지 브라우저 클릭 후 상단의 **저장소 탐색기로 다시 전환**을 클릭합니다.





1. **컨테이너**에서 마우스 오른쪽 클릭을 하여 **파일 시스템 만들기를 클릭**합니다.
2. **컨테이너를 추가**하기 위해 **이름을 입력**하고 확인을 클릭합니다. (4개의 컨테이너를 추가합니다.)



exportfile

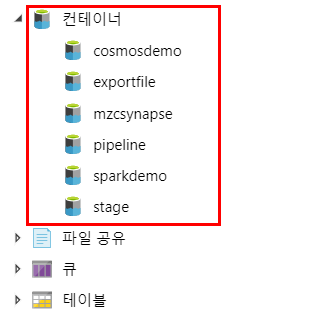
pipeline

stage

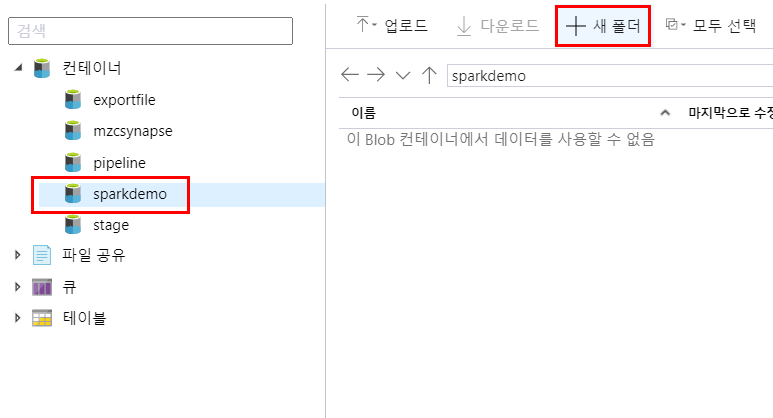
sparkdemo

cosmosdemo

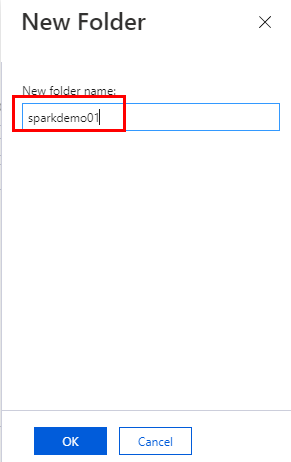
1. 총 **5개의 컨테이너가 구성**되었는지 확인합니다.



1. sparkdemo 컨테이너를 클릭합니다. **+새 폴더**를 클릭합니다.

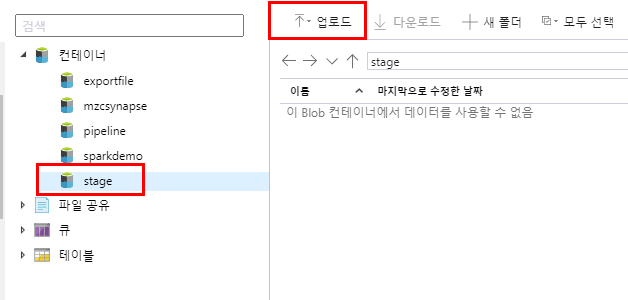


1. **sparkdemo01**이라는 폴더를 추가합니다.



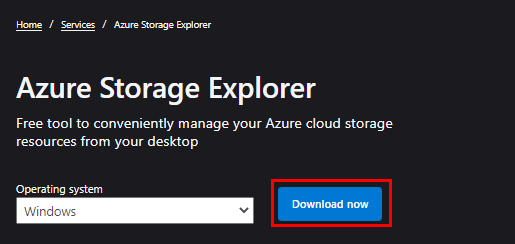
New folder name : sparkdemo01

1. 구성된 컨테이너 중에 하나를 선택하여 **업로드를 클릭**합니다.



1. 파일을 업로드하기 위해 **Azure Storage Explorer를 다운로드**하여 **설치**합니다.





텍스트이(가) 표시된 사진

자동 생성된 설명

1. 계정 연결을 위해 콘센트 모양의 버튼을 클릭합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

1. **리소스 선택**에서 **구독**을 클릭합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

1. **Azure 환경 선택**에서 **Azure 선택 후** **다음**을 클릭하여 **계정을 연결**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

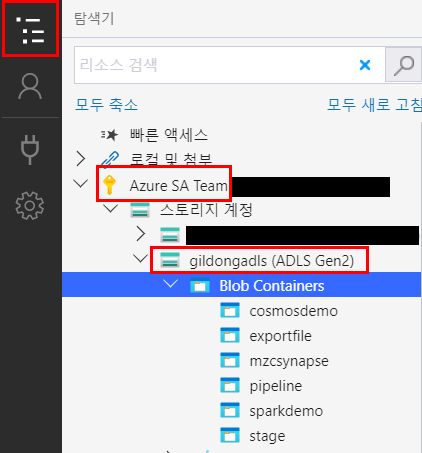
텍스트이(가) 표시된 사진

자동 생성된 설명

\* 연결이 완료되면 아래와 같은 문구가 나타납니다.

Authenticated. You can return to Storage Explorer. You might need to authenticate again if you close this browser tab.

1. 맨 위의 **왼쪽 상단을 클릭**하여 **개별 구독 아래에 스토리지 계정에서 앞에서 생성했던 ADLS Gen2**가 보이는지 확인합니다.



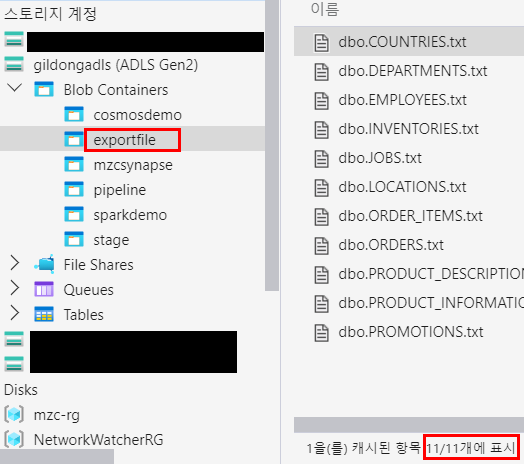
1. **Github**에서 실습에서 사용될 **자료를 다운**받습니다. 이미 자료을 받았다면 **압축을 풀어서 MZC Azure Synapse 폴더를 확인**합니다.

(<https://github.com/skagns88/mzc-azuresynapse>)

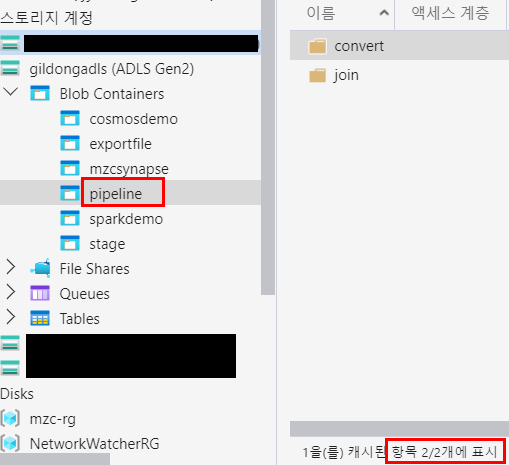
텍스트이(가) 표시된 사진

자동 생성된 설명

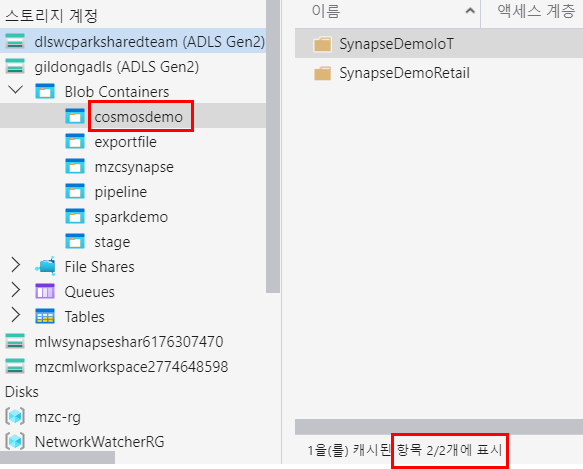
1. 다운로드한 자료에서 **MZC Azure Synapse\LAB File\exportfile안의 파일**들을 **exportfile 컨테이너**에 **Drag & Drop**하여 넣어줍니다. 11개의 텍스트 파일이 업로드 됩니다.



1. **MZC Azure Synapse\LAB File\pipeline안의 폴더**들을 **pipline 컨테이너**에 **Drag & Drop**하여 넣어줍니다. 2개의 폴더가 업로드 됩니다.



1. **MZC Azure Synapse\LAB File\cosmosdemo안의 폴더**들을 **cosmos 컨테이너**에 **Drag & Drop**하여 넣어줍니다. 2개의 폴더가 업로드 됩니다.



# Task 6 : Azure Synapse Analytics에서 SQL Pool 및 Spark 생성

1. 포털 상단 돋보기를 통해 **Azure Synapse Analytics를 검색**하여 **생성된 리소스를 클릭해서 이동**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

1. 왼쪽 블레이드에서 **SQL 풀로 이동**하여 **새로 만들기를 클릭**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

1. **기본 사항**에서 **SQL풀 세부 정보를 입력**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

전용 SQL 풀 이름 : sqlpool

성능 수준 : DW100c

1. **검토 + 만들기**를 클릭 후 하단의 **만들기**로 생성합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

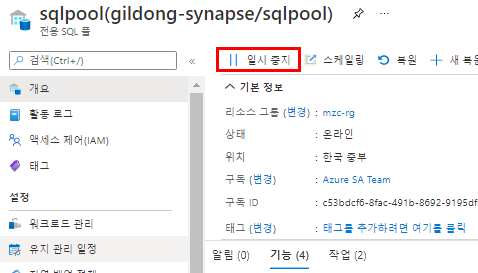
\* 그 외에 리소스 설정은 기본 값이므로 수정하지 마세요.

1. 리소스 생성이 완료되면 **다시 Azure Synapse Analytics 리소스로 돌아갑니다.** SQL 풀을 선택하여 **sqlpool이 보이는지 확인**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

\* sqlpool을 클릭하면 전용 SQL 풀에 대한 일시 중지가 가능합니다. 사용하지 않을 시 중지하여 비용 발생을 줄일 수 있습니다.



1. 왼쪽 블레이드에서 **Apache Spark 풀로 이동**하여 **새로 만들기를 클릭**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

1. **기본 사항**에서 **Apache Spark 풀 세부 정보를 입력**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

Apache Spark 풀 이름 : sparkpool

노드 크기 : Medium

자동 크기 조정 : 사용

노드 수 : 3 ~ 10

1. **검토 + 만들기**를 클릭 후 하단의 **만들기**로 생성합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

\* 그 외에 리소스 설정은 기본 값이므로 수정하지 마세요.

1. 리소스 생성이 완료되면 **다시 Azure Synapse Analytics 리소스로 돌아갑니다.** SQL 풀을 선택하여 **sparkpool이 보이는지 확인**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

# Task 7 : Azure Synapse Workspace Studio 접속

1. **Azure Synapse Analytics 리소스의 개요**에서 **작업 영역 웹 URL을 클릭**합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

1. **Azure Synapse Studio 웹 페이지가** 나타납니다.

텍스트이(가) 표시된 사진

자동 생성된 설명