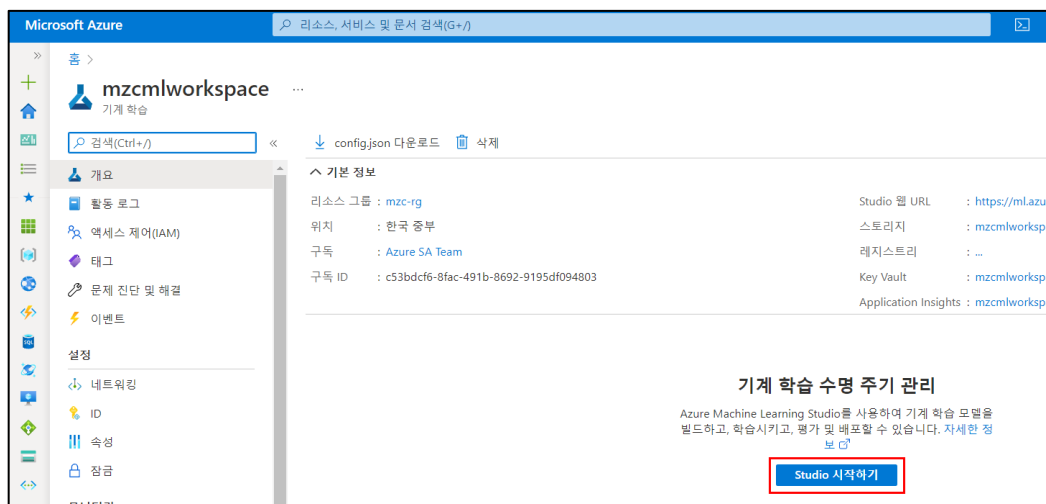


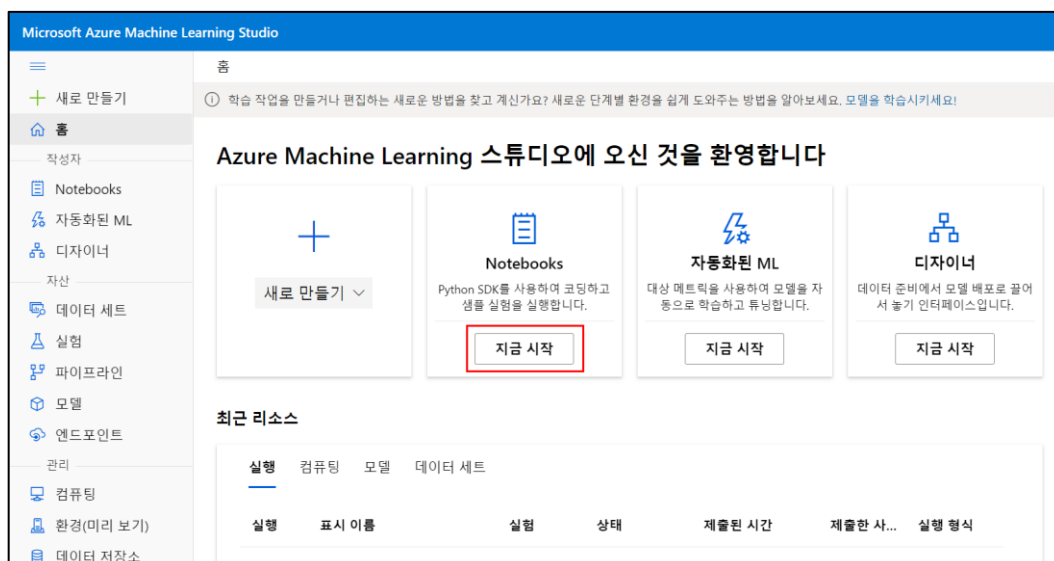
Lab 8 – Synapse SQL Pool 에서 Azure ML 모델 사용

Task 1 : Demo 05. Predict NYC Taxi Tips ONNX

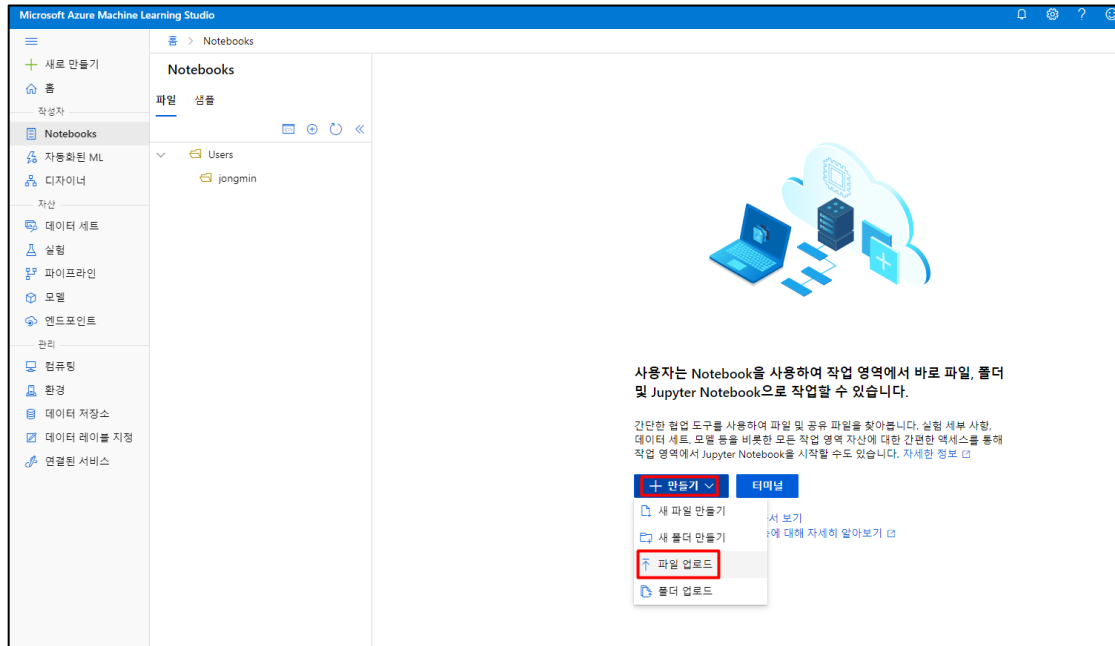
1. **Azure Machine Learning** 으로 접속합니다. **Launch Studio** 를 클릭하여 Studio 로 접속합니다.



2. Studio 에 접속하면 아래와 같은 페이지에 도달합니다. Notebooks 상자의 **지금 시작** 버튼을 클릭합니다.



3. 기존 Notebook 또는 파일 업로드를 클릭합니다.



4. Demo 05 번 파일을 아래와 같이 업로드합니다.

파일 업로드

파일 업로드 위치

Users/lyji [위치 편집](#)

+

파일을 찾아서 선택하려면 클릭하세요.

선택한 파일:

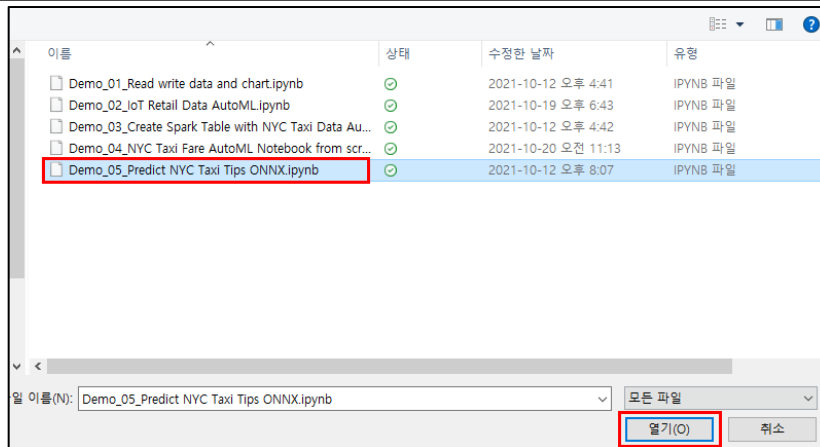
☐ 이미 있는 경우 덮어쓰기

☐ 이 파일의 내용을 신뢰함 *

로드하는 Notebook 또는 스크립트 내의 콘텐츠가 잠재적으로 세션에서 데이터를 읽고 Azure에서 조직 내의 데이터에 액세스할 수 있습니다. 소스 코드를 검토한 신뢰할 수 있는 출처의 Notebook 또는 스크립트만 Azure로 로드하세요. [Azure Machine Learning 관련 보안 코드 모범 사례에 대해 자세히 알아보세요.](#)

업로드

취소




5. 업로드를 눌러 업로드합니다.

파일 업로드

파일 업로드 위치

Users/jongmin [위치 편집](#)



선택한 파일을 바꾸려면 클릭하세요.

선택한 파일:

Demo_05_Predict NYC Taxi Tips ONNX.ipynb

☐ 이미 있는 경우 덮어쓰기

☒ 이 파일의 내용을 신뢰함 *

로드하는 Notebook 또는 스크립트 내의 콘텐츠가 잠재적으로 세션에서 데이터를 읽고 Azure에서 조직 내의 데이터에 액세스할 수 있습니다. 소스 코드를 검토한 신뢰할 수 있는 출처의 Notebook 또는 스크립트만 Azure로 로드하세요. [Azure Machine Learning 관련 보안 코드 모범 사례에 대해 자세히 알아보세요.](#)

업로드

취소

6. **컴퓨팅**으로 이동하여 만들어 놓은 컴퓨팅 인스턴스를 선택합니다. **시작** 버튼을 눌러 **실행중** 상태로 변경합니다.



컴퓨팅

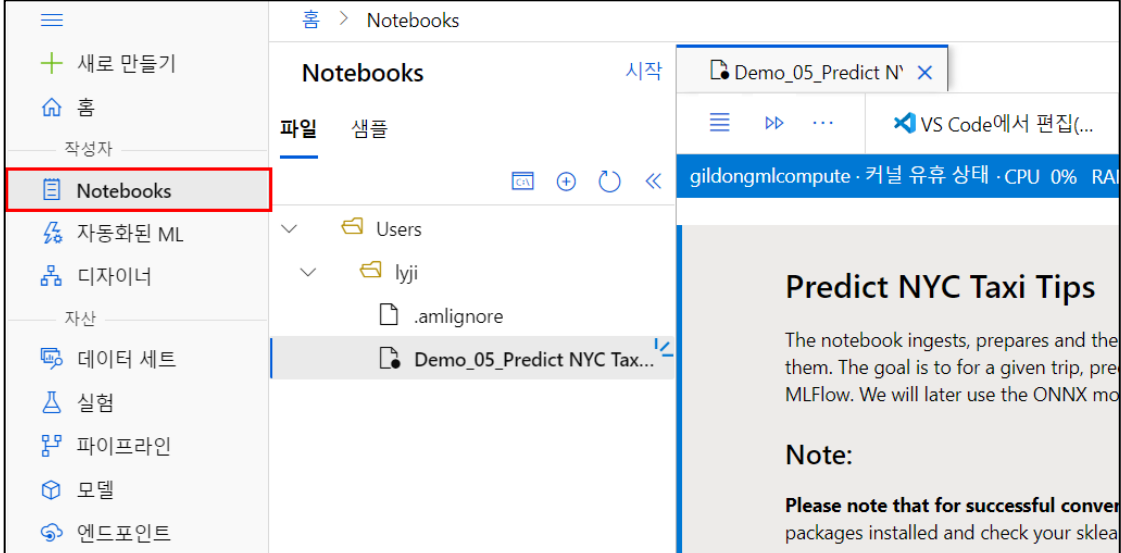
컴퓨팅 인스턴스 컴퓨팅 클러스터 유추 클러스터 연결된 컴퓨팅

+ 새로 만들기 새로 고침 **▶ 시작** 중지 다시 시작 삭제 열 편집 보기 다시 설정

검색

이름	상태	애플리케이션
gildongmlcompute	중지됨	JupyterLab Jupyter VS Code RStudio 터미널

7. 다시 **Notebooks** 로 이동합니다.



홈 > Notebooks

Notebooks 시작

파일 샘플

Users

lyji

.amlignore

Demo_05_Predict NYC Tax...

Demo_05_Predict N°

VS Code에서 편집(...)

gildongmlcompute · 커널 유휴 상태 · CPU 0% RAM

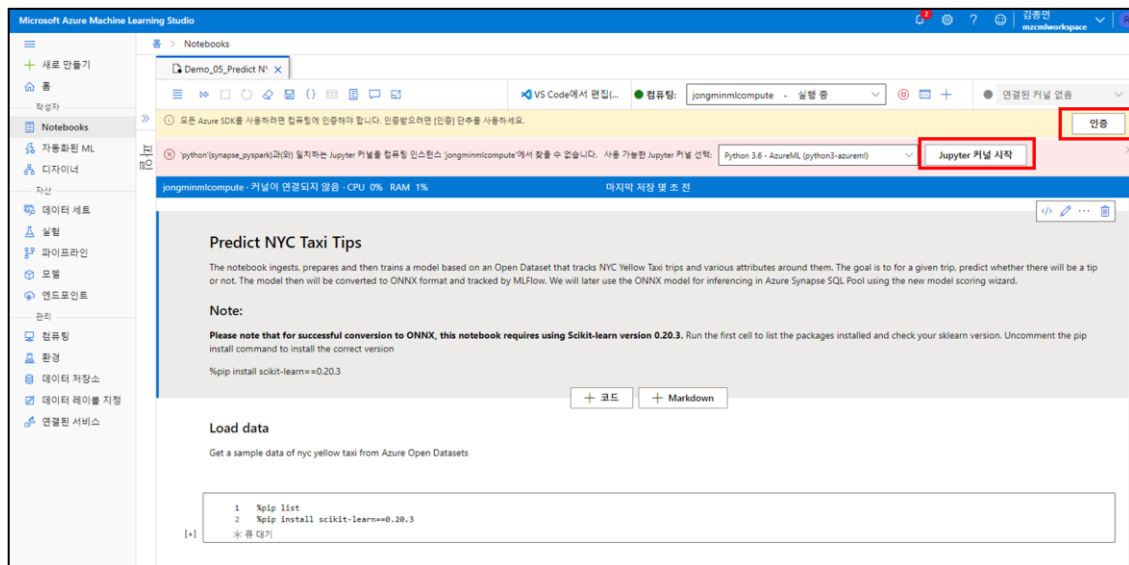
Predict NYC Taxi Tips

The notebook ingests, prepares and the them. The goal is to for a given trip, pre MLFlow. We will later use the ONNX mo

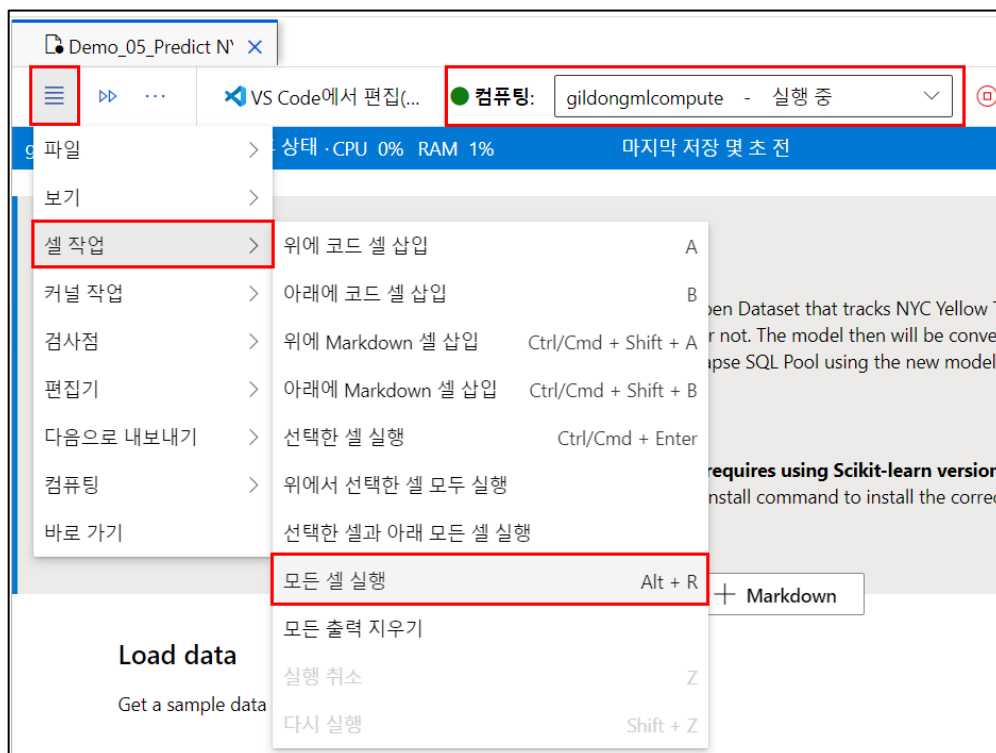
Note:

Please note that for successful conver packages installed and check your sklea

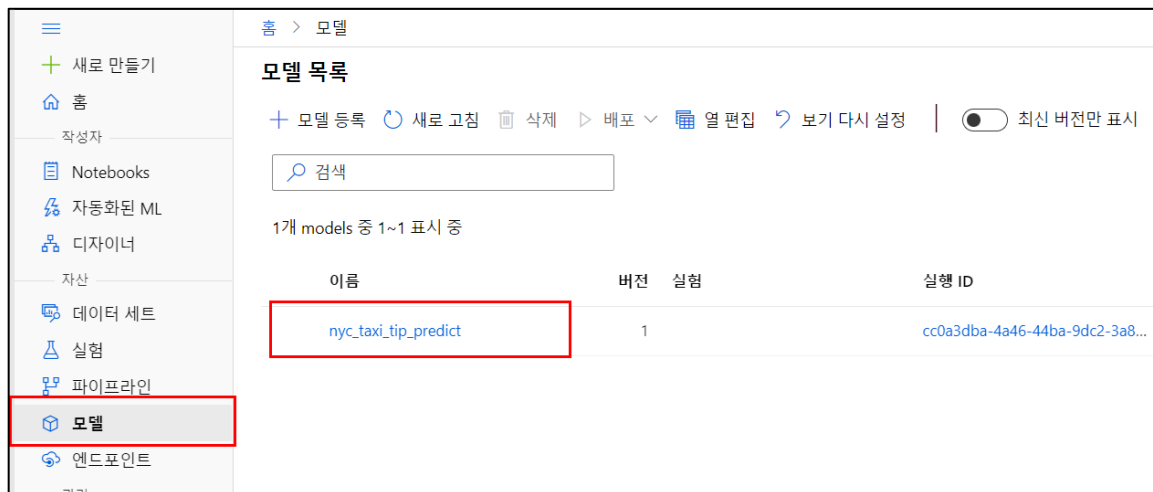
아래와 같이, 인증과 커널 시작을 눌러줍니다.



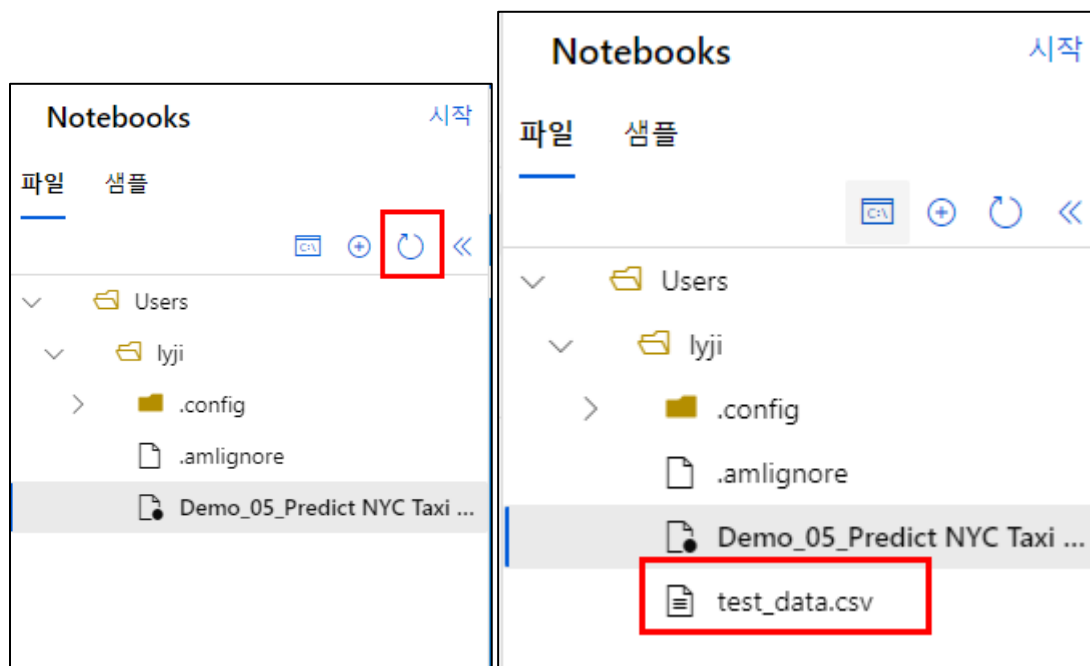
8. 컴퓨팅이 실행 중 상태가 된 것을 확인합니다. 메뉴 - 셀 작업 - 모든 셀 실행을 클릭합니다. 모든 셀이 실행되면서 모델이 만들어 집니다.



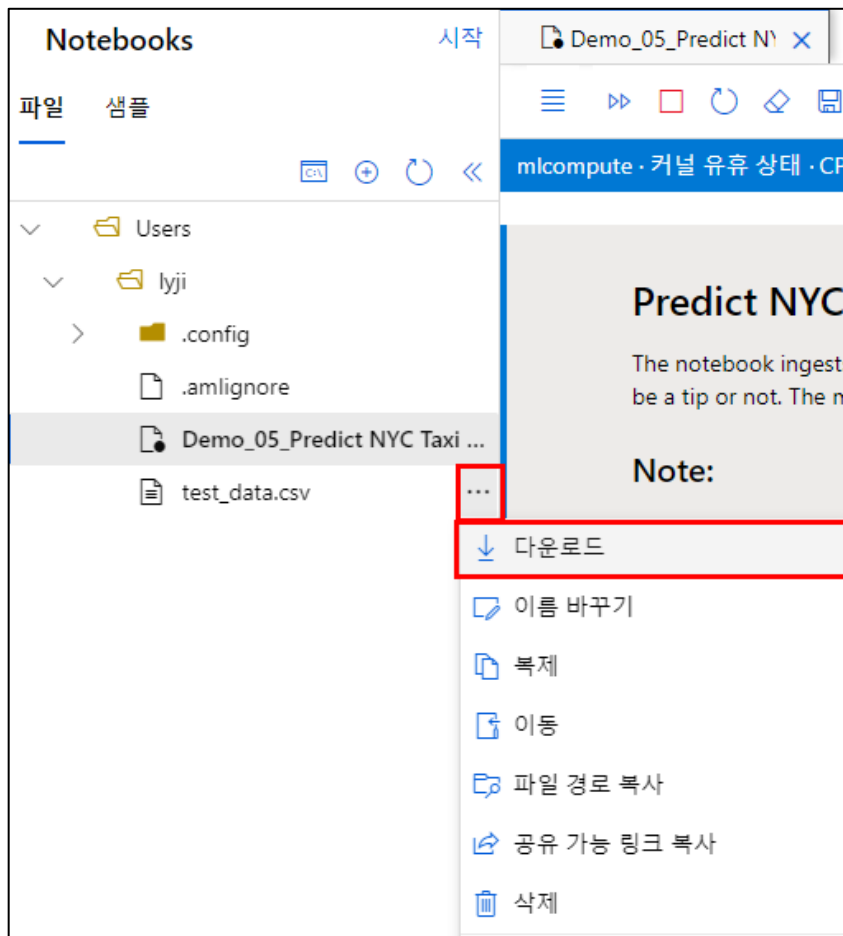
9. 실행이 완료되면 **모델** 로 이동합니다. 생성된 모델을 확인할 수 있습니다.



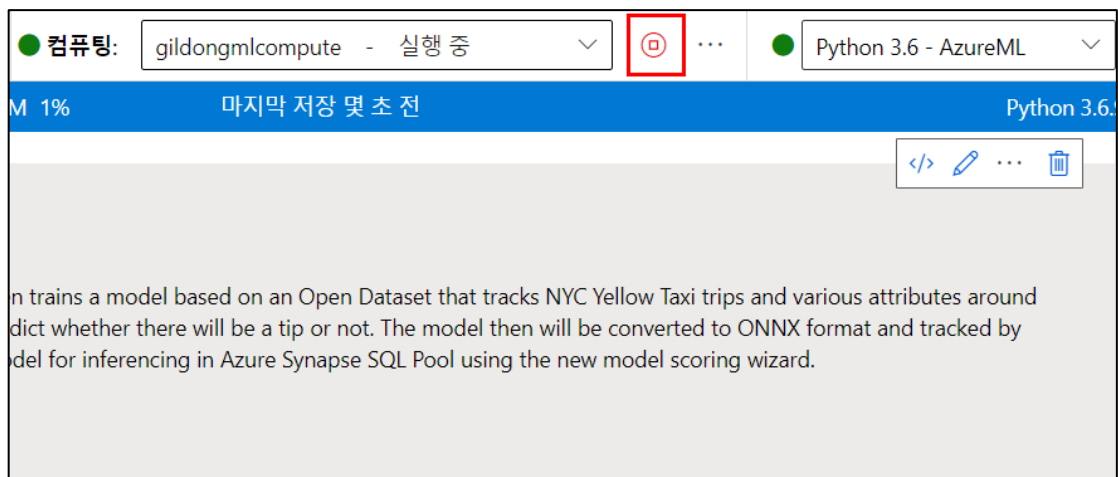
10. **Notebooks** 로 다시 이동합니다. **test_data.csv** 파일을 확인할 수 있습니다.
(보이지 않는 경우 파일 탭의 **새로고침** 버튼을 눌러봅니다.)



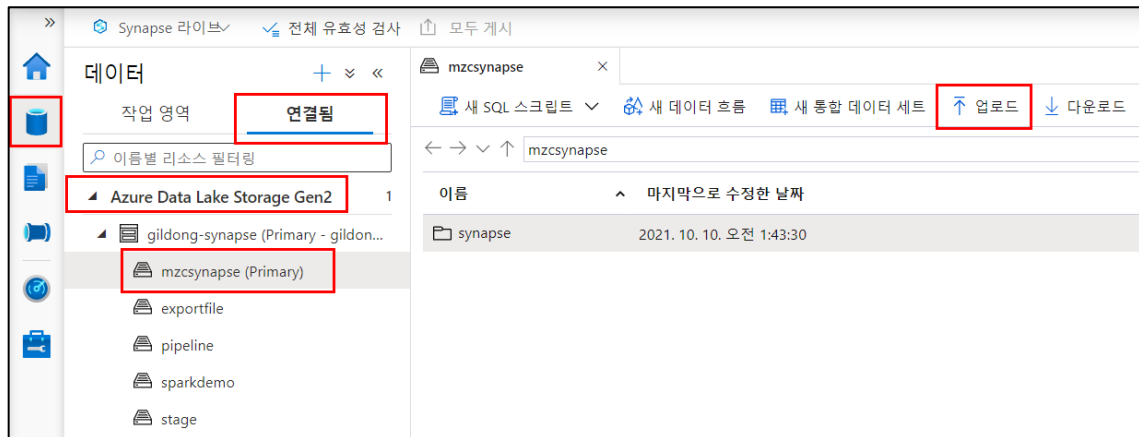
11. **test_data.csv** 파일을 다운로드 받습니다.



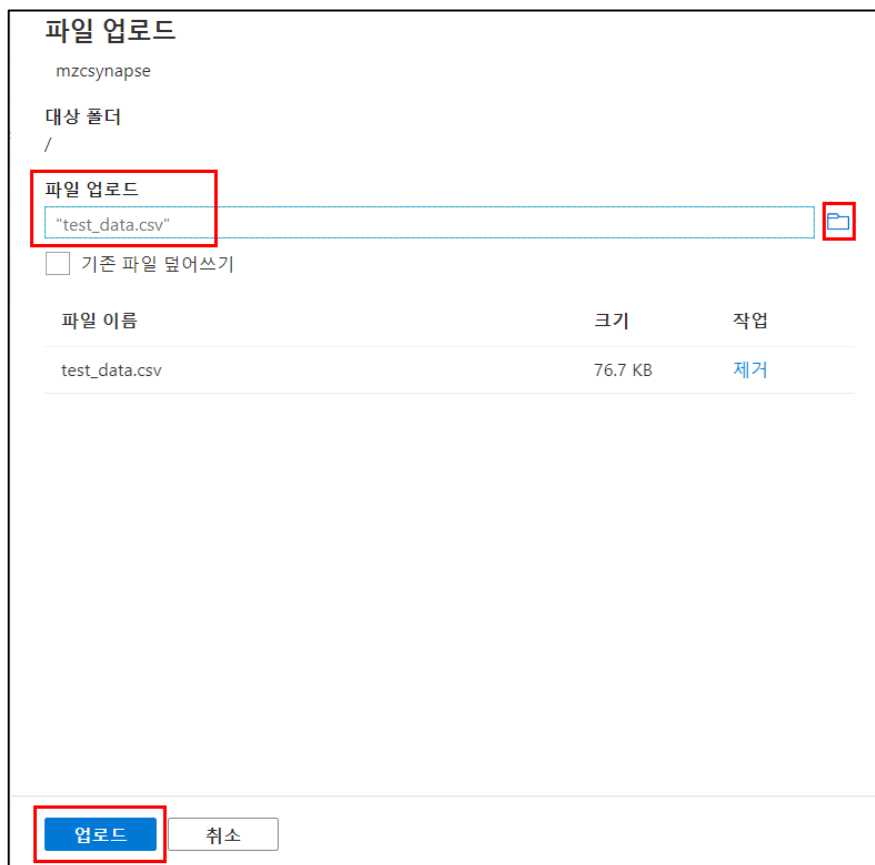
12. **컴퓨팅을 중지합니다.**



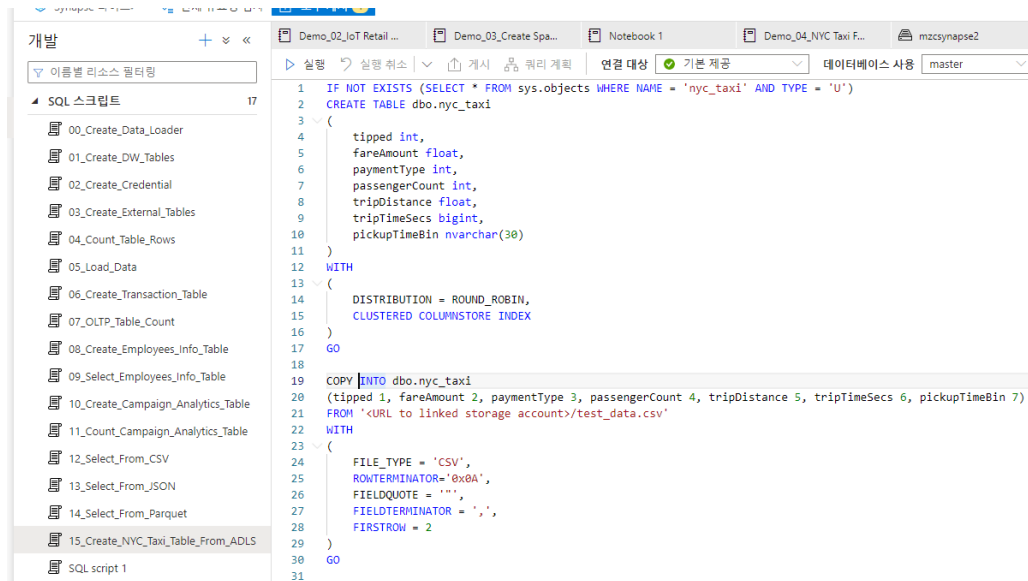
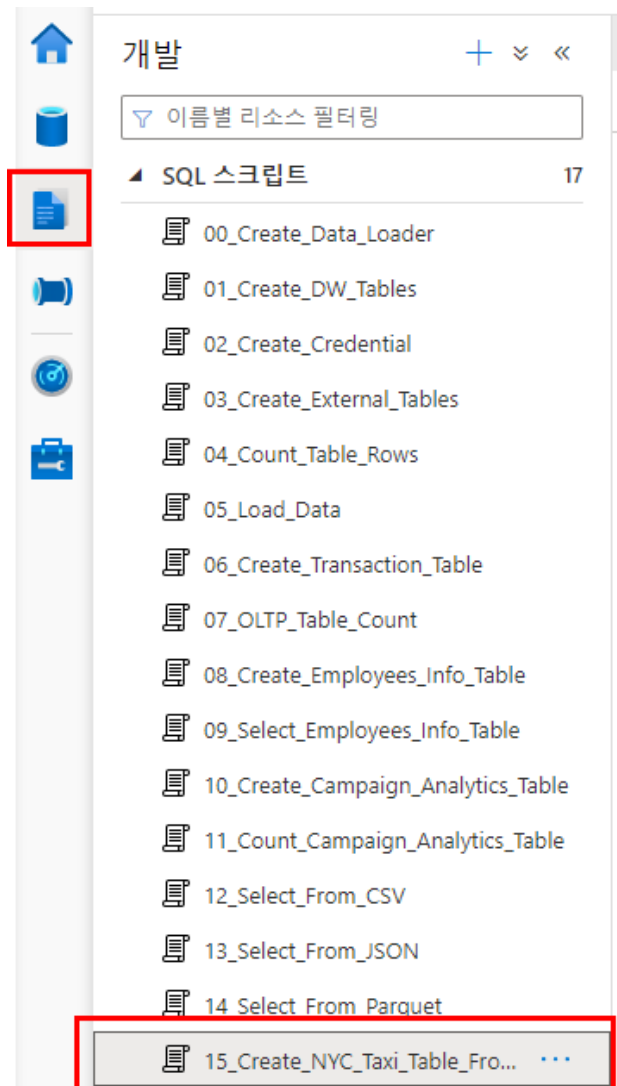
13. Synapse Studio 로 돌아옵니다. 데이터 허브에 접속하여 연결됨(Linked) 아래 ADLS 를 클릭합니다. mzcynapse 컨테이너를 클릭하고 업로드 버튼을 클릭합니다.



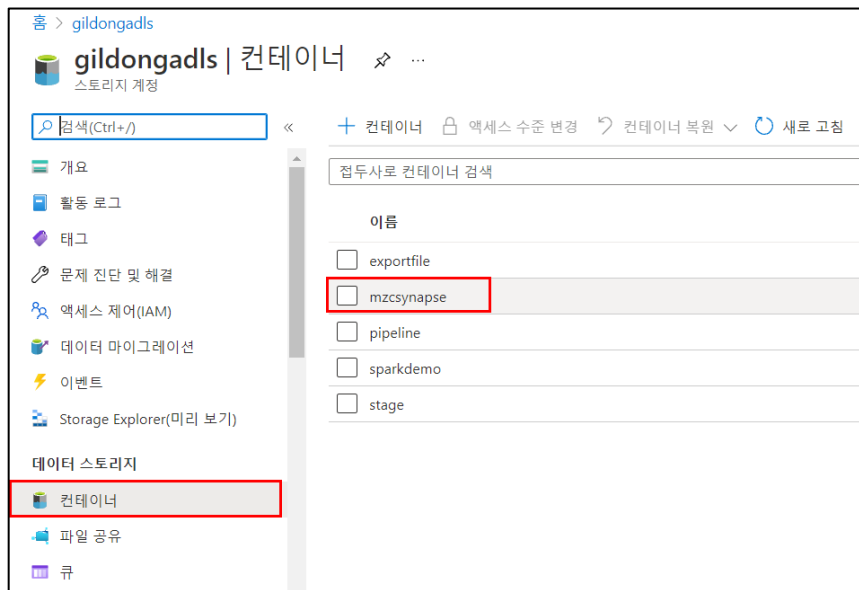
14. test_data.csv 파일을 선택하고 업로드합니다.



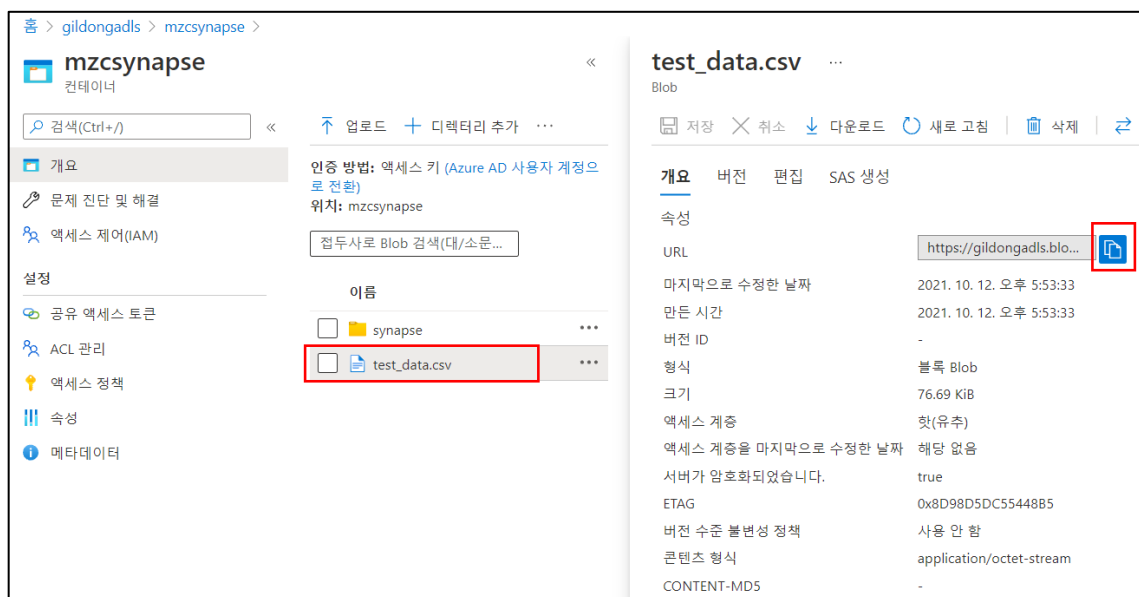
15. 개발 허브로 이동해서, SQL 스크립트 아래에 미리 import 해 놓은,
15_Create_NYC_Taxi_Table_from_ADLS 스크립트를 열어줍니다.



16. **ADLS 스토리지 계정**으로 이동합니다. 컨테이너를 선택하고 **test_data.csv** 파일이 있는 **mzcsynapse** 컨테이너로 이동합니다.



17. **test_data.csv** 파일을 클릭하고 URL 을 복사합니다.



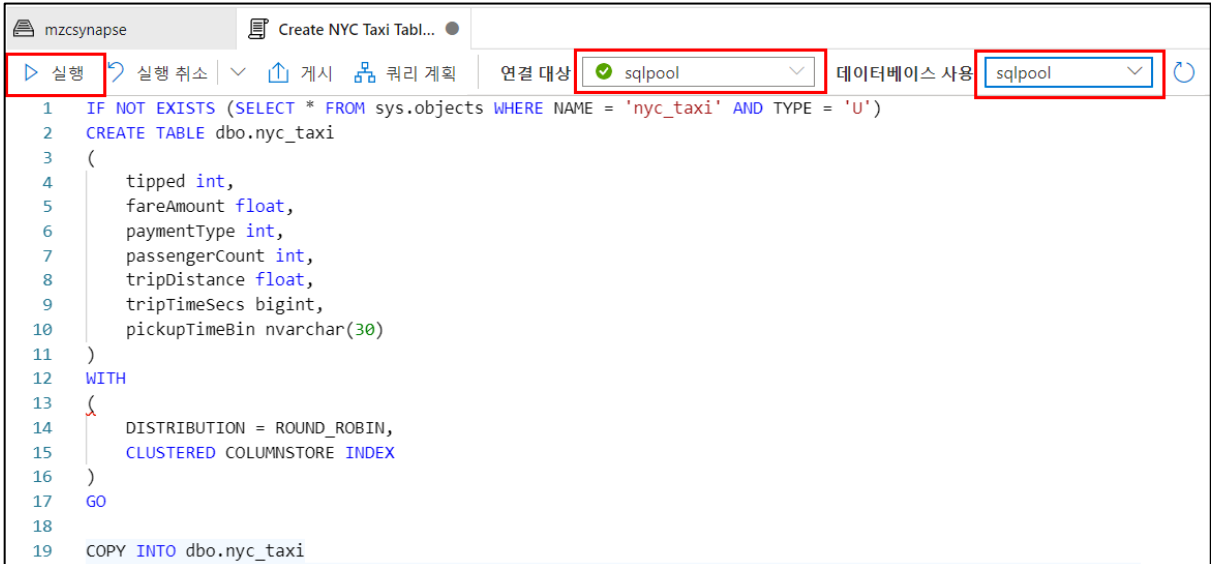
18. Synapse Studio 로 돌아옵니다. Create NYC Taxi Table from ADLS 스크립트의 해당 부분을 복사한 URL 값으로 변경합니다.

```

4      tipped int,
5      fareAmount float,
6      paymentType int,
7      passengerCount int,
8      tripDistance float,
9      tripTimeSecs bigint,
10     pickupTimeBin nvarchar(30)
11 )
12 WITH
13 (
14     DISTRIBUTION = ROUND_ROBIN,
15     CLUSTERED COLUMNSTORE INDEX
16 )
17 GO
18
19 COPY INTO dbo.nyc_taxi
20 (tipped 1, fareAmount 2, paymentType 3, passengerCount 4, tripDistance 5, tripTimeSecs 6, pickupTimeBin
21 FROM 'https://gildongadls.blob.core.windows.net/mzcsynapse/test_data.csv'
22 WITH
23 (
24     FILE_TYPE = 'CSV',
25     ROWTERMINATOR='0x0A',
26     FIELDQUOTE = '"',
27     FIELDTERMINATOR = ','

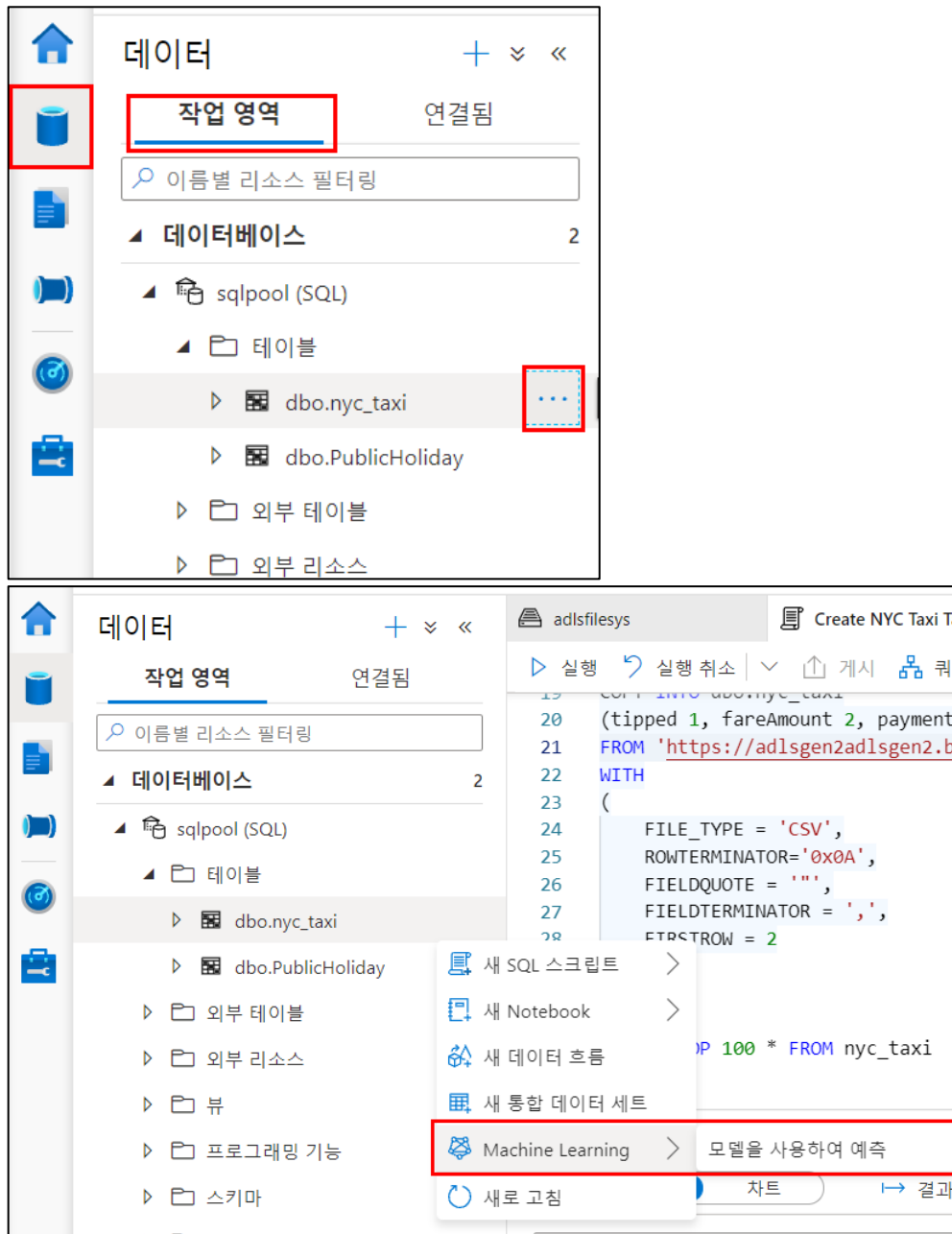
```

19. 연결대상으로 생성해놓은 전용 sql pool 을 선택합니다. 데이터베이스를 선택하고 실행버튼을 눌러 스크립트를 실행합니다.



The screenshot shows the Synapse Studio interface for the 'Create NYC Taxi Table' script. The 'Run' button (a blue play icon) is highlighted with a red box. The 'Connection' dropdown menu is set to 'sqlpool' and the 'Database' dropdown menu is also set to 'sqlpool', both highlighted with red boxes. The script content is visible in the background, showing the SQL code for creating the table and copying data from ADLS.

20. 데이터 허브로 이동합니다. 데이터베이스에서 **dbo.nyc_taxi** 테이블을 선택합니다. 오른쪽 **점 세 개**를 클릭하고 **Machine Learning – 모델을 사용하여 예측**을 선택합니다.



21. 만들어놓은 모델을 선택합니다. 계속을 클릭합니다.

모델을 사용하여 예측

dbo.nyc_taxi

선택한 데이터 세트를 보강하는 데 사용할 모델을 선택합니다. [자세한 정보](#)

Azure Machine Learning

Azure Machine Learning 작업 영역 *

mzcmllworkspace (AzureMLService)

이름	버전	생성 시간	만든 사람	프레...
nyc_taxi_tip_predict	1	08:37:15 10/1...	예지 이	Cust...

계속

취소

22. 입력, 출력 매핑을 설정(기본세팅을 유지)합니다. 계속을 클릭합니다.

모델을 사용하여 예측

dbo.nyc_taxi

원본 테이블 열을 필요한 모델 입력에 매핑합니다. [자세한 정보](#)

입력 매핑 *

+ 새로 만들기

삭제

<input type="checkbox"/> 원본 열	모델 입력	입력 형식		
<input type="checkbox"/> fareAmount	→ fareAmount	real	+	삭제
<input type="checkbox"/> paymentType	→ paymentType	bigint	+	삭제
<input type="checkbox"/> passengerCount	→ passengerCount	bigint	+	삭제
<input type="checkbox"/> tripDistance	→ tripDistance	real	+	삭제
<input type="checkbox"/> tripTimeSecs	→ tripTimeSecs	bigint	+	삭제
<input type="checkbox"/> pickupTimeBin	→ pickupTimeBin	varchar	+	삭제

출력 매핑 *

+ 새로 만들기

삭제

<input type="checkbox"/> 모델 출력	출력 형식		
<input type="checkbox"/> output_label	bigint	+	삭제

계속

뒤로

취소

MEGAZONE CLOUD Corp. All Right Reserved.

23. 저장 프로시저 이름과 대상 테이블을 설정합니다. 대상 테이블은 새로 만들기를 선택하고 이름을 설정합니다. 모델 배포 + 스크립트 열기를 클릭합니다.

모델을 사용하여 예측

dbo.nyc_taxi

저장 프로시저

생성된 스크립트를 실행하면 저장 프로시저가 생성됩니다. 이 저장 프로시저의 이름을 지정합니다. [자세한 정보](#)

저장 프로시저 이름 *

nyc_taxi_proc

대상 테이블

새 데이터베이스 테이블을 만들거나 기존 테이블을 사용하여 기계 학습 모델을 저장합니다. [자세한 정보](#)

대상 테이블 선택 *

☐ 기존 테이블

☒ 새로 만들기

새 테이블 *

nyc_taxi_aml_model

모델 배포 + 스크립트 열기

뒤로

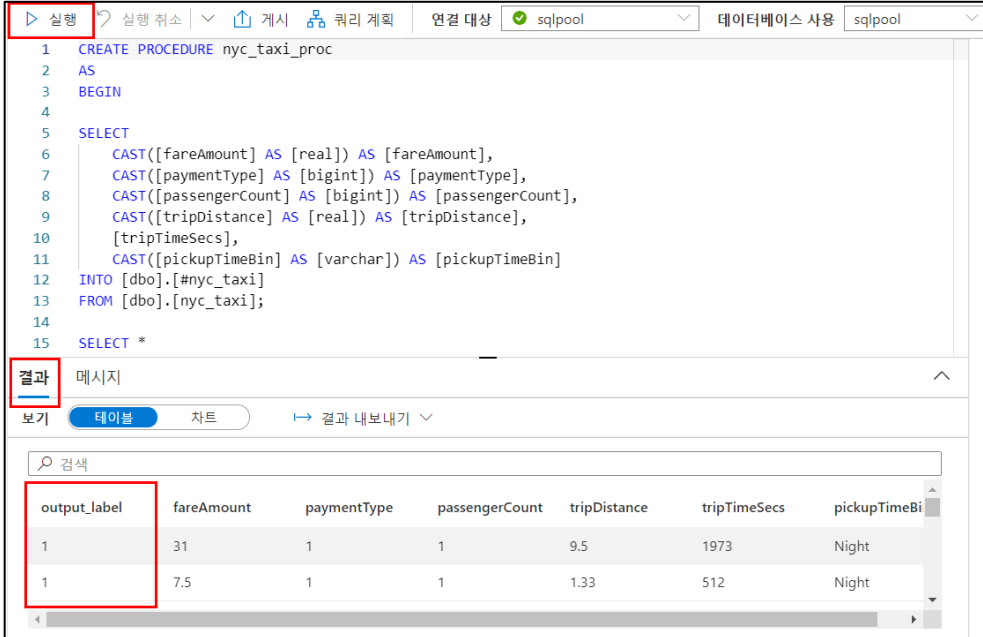
취소

저장 프로시저 이름: nyc_taxi_proc

대상 테이블 선택: 새로 만들기

새 테이블: nyc_taxi_aml_model

24. 만들어진 스크립트를 실행합니다. 결과를 확인합니다.



실행 실행 취소 | 게시 | 쿼리 계획 | 연결 대상: sqlpool | 데이터베이스 사용: sqlpool

```

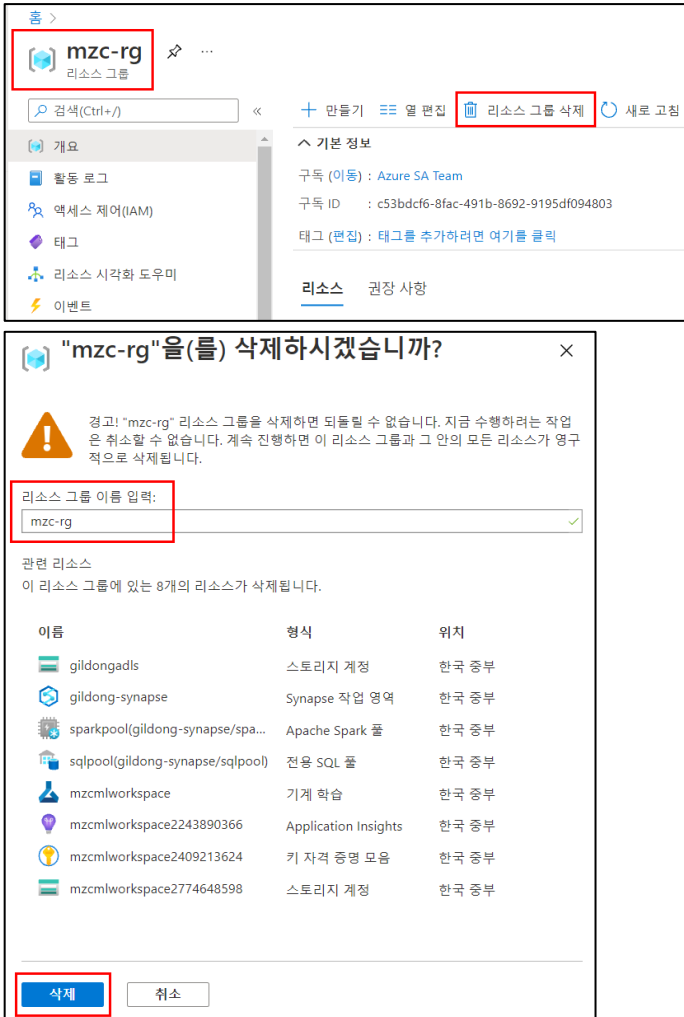
1 CREATE PROCEDURE nyc_taxi_proc
2 AS
3 BEGIN
4
5 SELECT
6     CAST([fareAmount] AS [real]) AS [fareAmount],
7     CAST([paymentType] AS [bigint]) AS [paymentType],
8     CAST([passengerCount] AS [bigint]) AS [passengerCount],
9     CAST([tripDistance] AS [real]) AS [tripDistance],
10    [tripTimeSecs],
11    CAST([pickupTimeBin] AS [varchar]) AS [pickupTimeBin]
12 INTO [dbo].[#nyc_taxi]
13 FROM [dbo].[nyc_taxi];
14
15 SELECT *
  
```

결과 메시지

보기: 테이블 | 차트 | 결과 내보내기

output_label	fareAmount	paymentType	passengerCount	tripDistance	tripTimeSecs	pickupTimeBin
1	31	1	1	9.5	1973	Night
1	7.5	1	1	1.33	512	Night

25. 실습이 끝난 후, 리소스 그룹을 삭제하여 추가적인 과금을 방지합니다.



mzc-rg 리소스 그룹

검색(Ctrl+/) | 만들기 | 열 편집 | **리소스 그룹 삭제** | 새로 고침

기본 정보

- 구독 (이동): Azure SA Team
- 구독 ID: c53bdcf6-8fac-491b-8692-9195df094803
- 태그 (편집): 태그를 추가하려면 여기를 클릭

리소스 | 권장 사항

"mzc-rg"을(를) 삭제하시겠습니까?

경고! "mzc-rg" 리소스 그룹을 삭제하면 되돌릴 수 없습니다. 지금 수행하려는 작업은 취소할 수 없습니다. 계속 진행하면 이 리소스 그룹과 그 안의 모든 리소스가 영구적으로 삭제됩니다.

리소스 그룹 이름 입력: mzc-rg

관련 리소스
이 리소스 그룹에 있는 8개의 리소스가 삭제됩니다.

이름	형식	위치
gildongadls	스토리지 계정	한국 중부
gildong-synapse	Synapse 작업 영역	한국 중부
sparkpool(gildong-synapse/spa...	Apache Spark 풀	한국 중부
sqlpool(gildong-synapse/sqlpool)	전용 SQL 풀	한국 중부
mzcmllworkspace	기계 학습	한국 중부
mzcmllworkspace2243890366	Application Insights	한국 중부
mzcmllworkspace2409213624	키 자격 증명 모음	한국 중부
mzcmllworkspace2774648598	스토리지 계정	한국 중부

삭제 | 취소