

한국 마이크로소프트

Microsoft Technical Trainer

Enterprise Skills Initiative

Microsoft Azure

Azure Stream Analytics

이 문서는 Microsoft Technical Trainer팀에서 ESI 교육 참석자분들에게 제공해 드리는 문서입니다.

요약

이 내용들은 표시된 날짜에 Microsoft에서 검토된 내용을 바탕으로 하고 있습니다. 따라서, 표기된 날짜 이후에 시장의 요구사항에 따라 달라질 수 있습니다. 이 문서는 고객에 대한 표기된 날짜 이후에 변화가 없다는 것을 보증하지 않습니다.

이 문서는 정보 제공을 목적으로 하며 어떠한 보증을 하지는 않습니다.

저작권에 관련된 법률을 준수하는 것은 고객의 역할이며, 이 문서를 마이크로소프트의 사전 동의 없이 어떤 형태(전자 문서, 물리적인 형태 막론하고) 어떠한 목적으로 재 생산, 저장 및 다시 전달하는 것은 허용되지 않습니다.

마이크로소프트는 이 문서에 들어있는 특허권, 상표, 저작권, 지적 재산권을 가집니다. 문서를 통해 명시적으로 허가된 경우가 아니면, 어떠한 경우에도 특허권, 상표, 저작권 및 지적 재산권은 다른 사용자에게 허여되지 않습니다.

© 2022 Microsoft Corporation All right reserved.

Microsoft®는 미합중국 및 여러 나라에 등록된 상표입니다.

이 문서에 기재된 실제 회사 이름 및 제품 이름은 각 소유자의 상표일 수 있습니다.

문서 작성 연혁

날짜	버전	작성자	변경 내용
2022.02.23	0.4.0	우진환	TASK 01 ~ TASK 03 작성
2022.02.24	0.8.0	우진환	TASK 04 ~ TASK 05 작성
2022.02.25	1.0.0	우진환	문서 편집 및 서식 정의
2022.04.04	1.1.0	우진환	Power BI 포털 UI 변경
2022.04.05	1.2.0	우진환	문서 서식 정의

목차

AZURE STREAM ANALYTICS 작동 방법.....	5
STREAM ANALYTICS의 INPUT과 OUTPUT.....	8
TASK 01. IOT 디바이스 시뮬레이터 구성.....	10
TASK 02. 입력(IOT HUB)과 출력(BLOB STORAGE) 구성	14
TAKS 03. STREAM ANALYTICS 작업 구성.....	17
TASK 04. BLOB STORAGE로 출력 구성.....	18
TASK 05. POWER BI로 집계 출력 구성.....	25
TASK 06. 리소스 정리.....	40

Azure Stream Analytics는 여러 소스에서 대량의 스트리밍 데이터를 동시에 분석하고 처리하도록 디자인된 실시간 분석/복잡한 이벤트 처리 엔진입니다. 패턴과 관계는 디바이스, 센서, 클릭스트림(clickstream), 소셜 미디어 피드, 애플리케이션을 포함한 다양한 입력 소스에서 추출한 정보로부터 식별할 수 있습니다. 이러한 패턴은 작업을 트리거하고 워크플로를 시작하는데 사용할 수 있습니다. 예를 들어 경고를 생성하거나 리포팅 도구로 정보를 제공하거나 추후 사용을 위해 변환된 데이터를 저장할 수 있습니다. 또한 Stream Analytics는 Azure IoT Edge 런타임에서 사용할 수 있으므로 IoT 디바이스의 데이터를 처리할 수 있습니다. 다음과 같은 시나리오에서 Azure Stream Analytics를 사용할 수 있습니다.

- IoT 디바이스에서 실시간 원격 분석(real-time telemetry) 스트림 분석
- 웹 로그/클릭스트림 분석
- 차량 관리 및 무인 차량을 위한 지리 공간 분석
- 높은 가치의 자산에 대한 원격 모니터링 및 예측적 유지 보수
- 재고 관리 및 이상 탐지를 위한 PoS (Point of Sale) 데이터의 실시간 분석

Azure Stream Analytics는 다음과 같은 주요 특징이 있습니다.

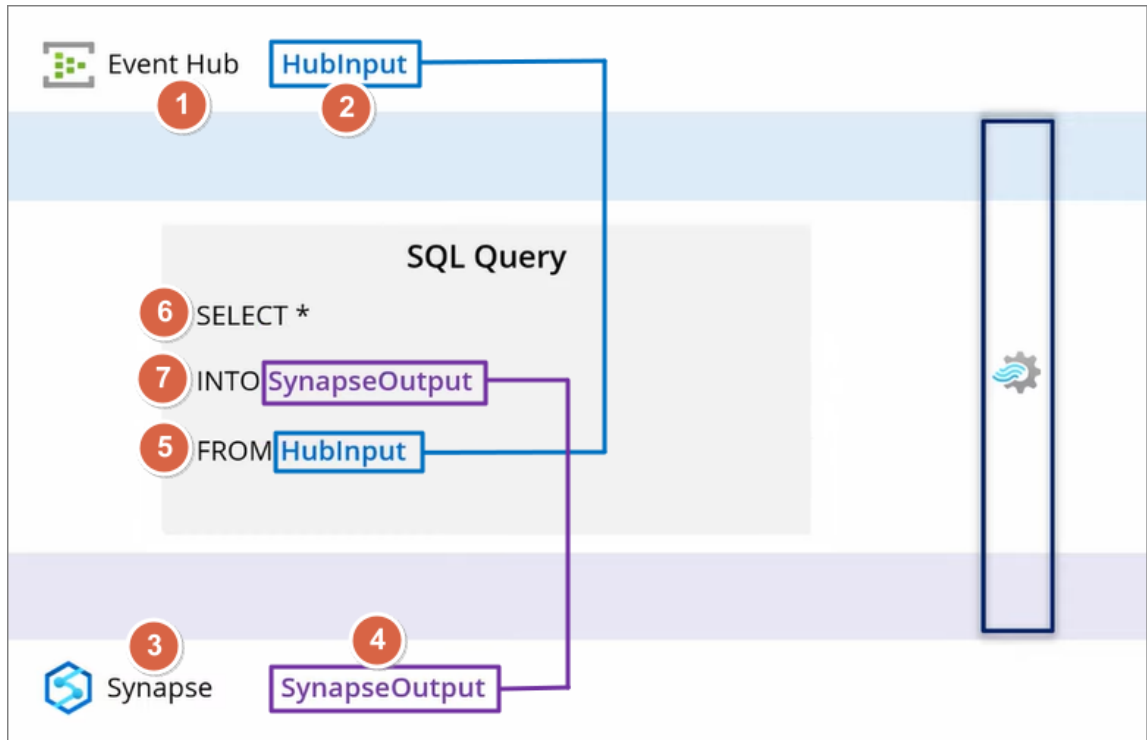
- 친숙한 SQL 쿼리 언어를 사용하여 데이터 스트림을 분석할 수 있습니다. SQL은 많은 데이터 엔지니어들이 데이터 분석을 사용하는 일반적인 언어이기 때문에 Stream Analytics와 관련된 다양한 참조를 쉽게 찾을 수 있습니다.
- 여러 통합으로 다른 Azure 서비스에 대한 out-of-box 커넥터를 제공하며 기존 인프라에 Stream Analytics를 쉽게 구현할 수 있습니다.
- SQL 쿼리 언어로 스트리밍 데이터에 대한 실시간 분석 논리를 구현할 수 있지만 SQL 쿼리 언어로 충분하지 않은 경우 쿼리에서 호출되는 사용자 지정 함수(custom function)를 통해 추가적인 유연성을 제공합니다. 이를 위해 JavaScript, C#, Azure Machine Learning을 사용할 수 있습니다.

Azure Stream Analytics 작동 방법

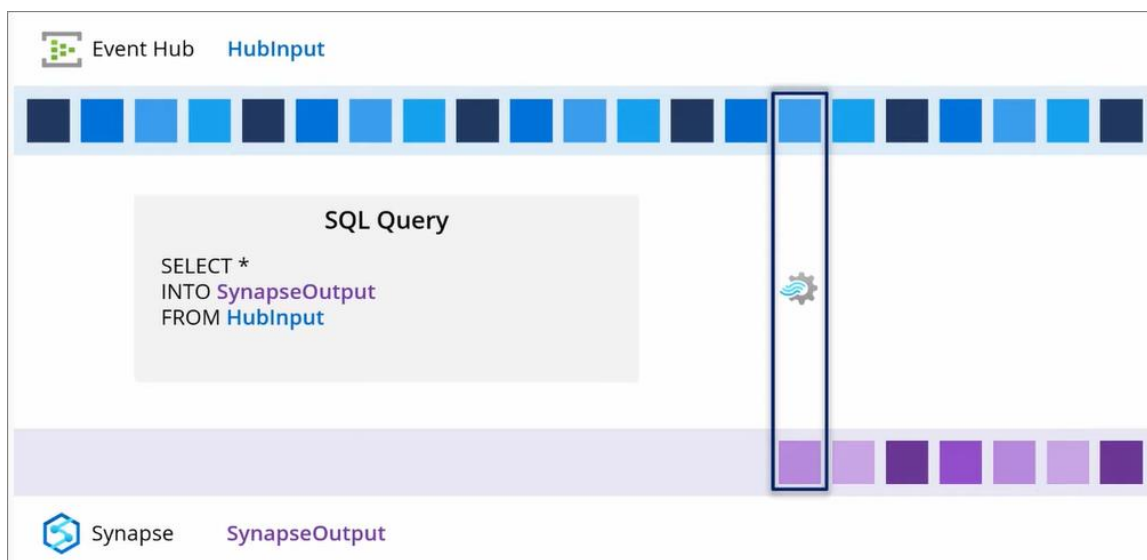
Azure Stream Analytics는 다음과 같은 방법으로 작동합니다.

- ① 소스의 유형과 속성을 정의하고 몇 가지 공통적인 속성을 입력하여 어떤 소스에서 시작할 것인지 정의합니다.
- ② 소스에 대한 별칭(alias)을 지정합니다. 아래 이미지에서는 Azure Event Hub가 소스이며 이에 대한 별칭으로 **HubInput**을 사용합니다.
- ③ 입력 소스가 정의되면 출력을 정의합니다. 아래 이미지에서는 Synapse Analytics가 출력이 됩니다.
- ④ 출력에 대한 별칭을 정의합니다. 아래 이미지에서는 **SynapseOutput**을 사용합니다.
- 입력과 출력이 모두 정의되면 SQL 작업 쿼리를 시작할 수 있습니다. 이 쿼리는 Stream Analytics가 데이터를 처리하는 핵심입니다. 작업 쿼리에서 데이터를 선택하는 방법, 적용해야 하는 변환 유형, 출력하는 방법에 대해 정의합니다.
- ⑤ 표준 SQL과 마찬가지로 **FROM** 문에서 사용할 입력 소스의 별칭을 입력하여 소스에서 데이터를 가져옵니다.

- ⑥ **SELECT** 문을 사용하여 변환을 수행합니다. 아래 이미지에서는 모든 내용을 가져옵니다.
- ⑦ **INTO** 문에 출력에 대한 별칭을 사용하여 변환한 내용을 출력할 대상을 지정합니다.

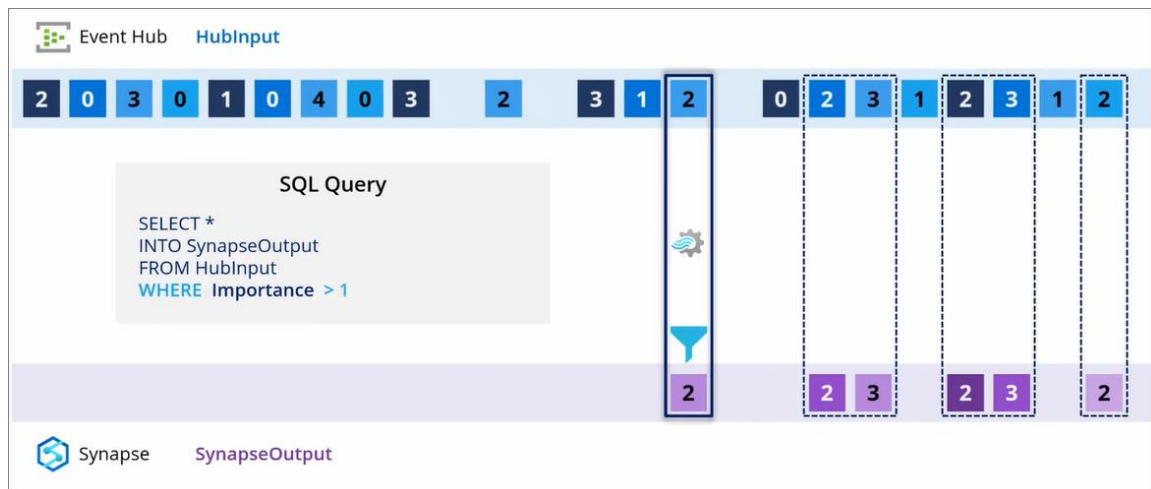


Stream Analytics에 대한 구성이 완료되면 작업을 시작할 수 있습니다. 작업이 시작되면 소스에서 데이터가 들어오고 Stream Analytics 작업의 출력으로 데이터가 복사됩니다. 아래 이미지의 경우 **SELECT ***를 사용했기 때문에 전체 입력 스트림을 출력에 모두 복사합니다. 즉 Event Hub의 모든 데이터를 가져와 Synapse Analytics로 데이터를 스트리밍하게 됩니다.



Stream Analytics는 일반 SQL과 마찬가지로 필터링을 적용할 수 있습니다. 예를 들어 Event Hub의 입력 데이터 스키마가 **ID**, **Importance**, **User** 이름의 열로 구성되어 있다고 가정해 보겠습니다. 이 경우 SQL 쿼리 내의 **WHERE** 문에서 **Importance**의 값이 **1**보다 큰 값만 출력으로 변환되도록 설정할 수 있습니다. 이렇게

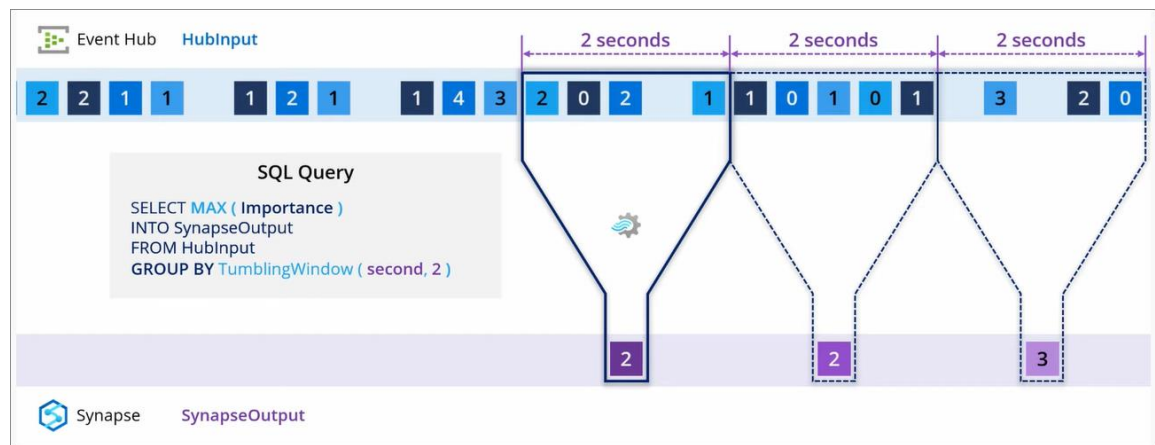
구성하게 되면 Event Hub에서 들어온 데이터는 해당 조건을 기반으로 출력을 필터링하게 됩니다.



또한 특정 분석을 위해 일정 시간을 기반으로 작업을 수행할 수 있습니다. 이를 위해 일반 SQL과 마찬가지로 **GROUP BY** 문을 사용하고 시간 간격을 정의하는 **TumblingWindow** 함수를 사용할 수 있습니다. 아래 예의 경우 2초마다 수신된 데이터의 **Importance** 열의 최대 값인 **MAX(Importance)**를 계산합니다. **GROUP BY**를 사용하기 때문에 일반 SQL과 마찬가지로 집계 함수(**MAX**, **MIN**, **AVG** 등)를 사용해야 합니다. 이렇게 구성되면 수신된 데이터는 2초마다 그룹화되고 결과가 계산된 후 결과가 출력으로 복사됩니다.

시간 스트리밍(time-streaming) 시나리오에서 임시 기간(temporal window)에 포함된 데이터에 대한 작업을 수행하는 것은 매우 일반적인 패턴입니다. Stream Analytics는 **windowing** 함수를 기본적으로 지원하기 때문에 복잡한 스트림 처리를 쉽게 작성할 수 있습니다. 사용할 수 있는 임시 기간(temporal window)은 **Tumble**, **Hopping**, **Sliding**, **Session**, **Snapshot**의 5가지입니다.

- **TumblingWindow**는 데이터 스트림을 별개의 시간 세그먼트로 분할합니다. **TumblingWindow**는 반복되고 겹치지 않으며, 이벤트가 둘 이상의 **TumblingWindow**에 포함될 수 없습니다.
- **HoppingWindow**는 고정된 시간만큼 앞으로 **hop** 합니다. 이는 **window size**가 겹칠 수 있고 더 자주 방출될 수 있는 **TumblingWindow**로 생각할 수 있습니다. 따라서 이벤트는 둘 이상의 **HoppingWindow** 집합에 속할 수 있습니다. **HoppingWindow**를 **TumblingWindow**와 동일하게 구성하려면 **hop size**를 **window size**와 동일하게 지정하면 됩니다.
- **SlidingWindow**는 **TumblingWindow**, **HoppingWindow**와 달리 window의 콘텐츠가 실제로 변경되는 특정 시점에 대해서만 이벤트를 출력합니다. 예를 들어 지난 10초 동안 3번 이상 언급된 토픽이 있는 경우 알림을 받고자 하는 경우 이를 사용할 수 있습니다. 따라서 모든 window에는 하나 이상의 이벤트가 있습니다.
- **SessionWindow** 함수는 유사한 시간에 도착하는 이벤트를 그룹화하여 데이터가 없는 기간을 필터링합니다. 예를 들어 서로 5초 이내에 발생한 트윗 수를 확인할 때 이 함수를 사용할 수 있습니다.
- **Snapshot window**는 타임스탬프가 동일한 이벤트를 그룹화합니다. 특정 window 함수(예를 들어 **SessionWindow()**)가 필요한 다른 windowing 유형과 달리 **Snapshot window**는 **GROUP BY** 절에 **System.Timestamp()**를 추가하여 적용할 수 있습니다. 예를 들어 정확한 시간에 발생한 동일한 토픽 유형의 트윗 수를 확인하고자 하는 경우 이 함수를 사용할 수 있습니다.



Stream Analytics의 input과 output

Azure Stream Analytics 작업은 하나 이상의 데이터 입력에 연결합니다. 각 입력은 기존 데이터 소스에 대한 연결을 정의합니다. Stream Analytics는 Azure Event Hubs, Azure IoT Hub, Azure Blob Storage, Azure Data Lake Storage Gen2에서 들어오는 데이터를 허용합니다. 데이터가 데이터 원본으로 푸시되면 Stream Analytics 작업에서 사용되고 실시간으로 처리됩니다. 입력은 데이터 스트림 입력과 참조 데이터 입력의 두 가지 유형으로 구분됩니다.

- 데이터 스트림(data stream)은 시간 경과에 따라 계속 발생하는 무한한 이벤트 시퀀스입니다. Stream Analytics 작업에는 데이터 스트림 입력이 하나 이상 포함되어야 하며 Event Hub, IoT Hub, Blob Storage, Data Lake Gen2가 데이터 스트림의 입력 소스로 지원됩니다. Event Hubs는 여러 디바이스 및 서비스에서 이벤트 스트림을 수집하는데 사용됩니다. 이러한 스트림에는 소셜 미디어 활동 피드, 주식 거래 정보, 센서의 데이터가 포함될 수 있습니다. IoT Hub는 IoT 시나리오에서 연결된 디바이스의 데이터를 수집하도록 최적화되어 있습니다. Blob Storage는 로그 파일과 같은 스트림으로 대량 데이터를 수집하기 위한 입력 소스로 사용할 수 있습니다.
- 참조 데이터(reference data)는 정적 데이터 혹은 느리게 변경되는 데이터입니다. 이 데이터는 상관 관계 및 조회를 수행하는데 사용됩니다. 예를 들어 정적 값을 조회하기 위해 SQL 조인을 수행하는 것처럼 데이터 스트림 입력의 데이터를 참조 데이터의 데이터에 조인할 수 있습니다. 예를 들어 스트림 데이터와 결합하고 추가 분석을 수행하기 위해 애플리케이션의 일부 지리적 데이터나 사전 데이터를 참조 데이터로 설정하여 사용할 수 있습니다.

Azure Stream Analytics의 출력을 위해 지원되는 많은 서비스가 있습니다. Stream Analytics 쿼리를 디자인할 때 여러 **INTO** 절을 사용하여 출력 이름을 참조할 수 있으며 쿼리에 여러 **INTO** 절을 추가하여 작업에 여러 출력을 사용할 수 있습니다. 따라서 단일 작업과 단일 쿼리를 사용하여 데이터를 결합하거나 데이터를 여러 출력으로 출력할 수 있습니다.

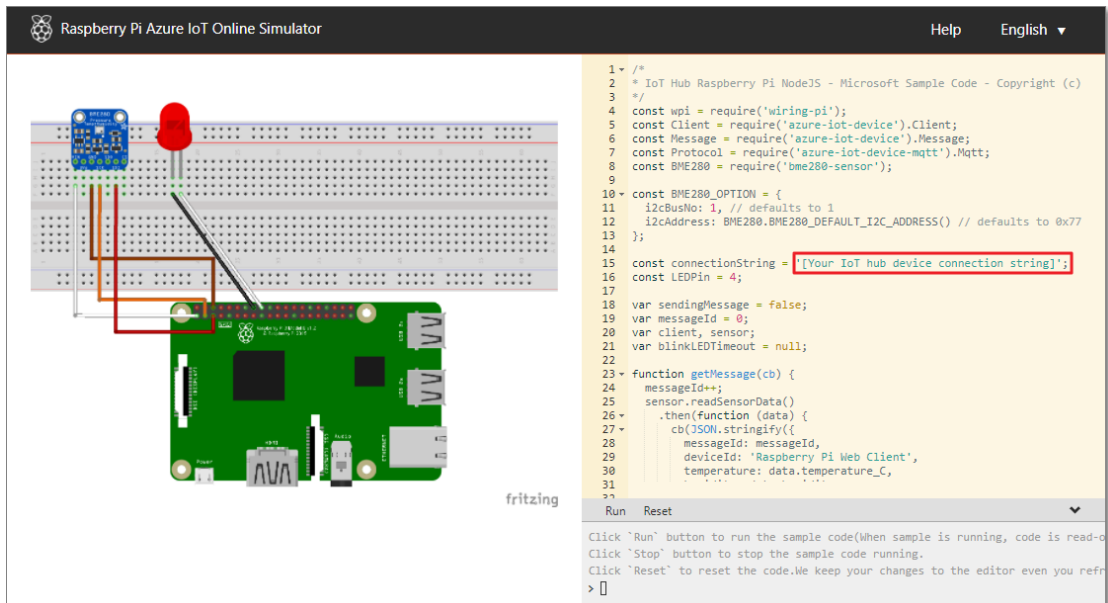
- 출력에 일반적으로 Azure Blob, Azure Data Lake, Event Hub, Service Bus, Synapse Analytics, Functions, Power BI 등을 사용할 수 있습니다.
- Power BI로의 데이터 스트리밍은 Azure Stream Analytics의 핵심 사용 사례 중 하나입니다. Power BI에서

스트림에 대해 쿼리를 적용하면 라이브 대시보드를 볼 수 있습니다.

TASK 01. IoT 디바이스 시뮬레이터 구성

이 실습에서는 브라우저 기반 Raspberry Pi 시뮬레이터를 사용합니다.

1. 브라우저를 열고 [Raspberry Pi Azure IoT Web Simulator \(azure-samples.github.io\)](https://azure-samples.github.io/RaspberryPiAzureIoTWebSimulator/) 사이트로 이동합니다. 아래와 같이 Raspberry Pi 웹 시뮬레이터가 표시됩니다. 우측에 해당 시뮬레이터에 대한 코드를 찾을 수 있습니다. 코드 부분을 확인하면 `connectionString`이 있고 이를 IoT Hub 연결 문자열로 교체하면 이 시뮬레이터는 해당 IoT Hub로 데이터를 전송하게 됩니다.



2. IoT 시뮬레이터를 실행하였으므로 이제 IoT Hub를 만들어야 합니다. IoT 디바이스는 IoT Hub로 데이터를 전송하고 Stream Analytics에서 IoT Hub를 입력으로 사용하게 됩니다.
3. Azure 포털에서 [리소스 만들기]를 클릭한 후 "IoT Hub"를 검색하고 클릭합니다. [IoT Hub] 블레이드에서 [만들기]를 클릭합니다.



4. [IoT 허브] 블레이드의 [기본 사항] 탭에서 아래와 같이 구성한 후 [다음]을 클릭합니다.
 - [프로젝트 정보 - 리소스 그룹]: "새로 만들기"를 클릭한 후 "05_streamAnalyticsRg"를 입력합니다.
 - [인스턴스 정보 - IoT Hub 이름]: 중복되지 않는 고유한 이름을 입력합니다.
 - [인스턴스 정보 - 지역]: Korea Central

IoT 허브 ...

Microsoft

기본 사항 네트워크 관리 태그 검토 + 만들기

IoT 허브를 만들면 수많은 IoT 자산을 연결, 모니터링 및 관리하는 데 도움이 됩니다. [자세한 정보](#)

프로젝트 정보

배포 및 비용을 관리하는 데 사용할 구독을 선택합니다. 폴더와 같은 리소스 그룹을 사용하면 리소스를 구성하고 관리하는 데 도움이 됩니다.

구독 * ⓘ Azure Pass - 스폰서쉽

리소스 그룹 * ⓘ (신규) 05_streamAnalyticsRg 새로 만들기

인스턴스 정보

IoT Hub 이름 * ⓘ kormttiotHub

지역 * ⓘ Korea Central

5. [네트워크] 탭에서 기본값을 유지하고 [다음]을 클릭합니다.

IoT 허브 ...

Microsoft

기본 사항 **네트워크** 관리 태그 검토 + 만들기

공용 호스트 이름을 통해 공개적으로 또는 프라이빗 엔드포인트를 사용하여 비공개로 IoT Hub에 연결할 수 있습니다. [자세히 알아보기](#)

연결 구성 *

☒ 공용 액세스

☐ 프라이빗 액세스(권장)

이 리소스가 생성된 후 이를 변경하거나 다른 연결 방법을 구성할 수 있습니다. [자세히 알아보기](#)

6. [관리] 탭에서 가격 및 크기 계층을 "S1: 표준 계층"으로 선택하고 "Defender for IoT" 설정은 해제합니다. 다른 값은 기본값을 유지한 후 [검토 + 만들기]를 클릭합니다. [검토 + 만들기] 탭에서 [만들기]를 클릭합니다.

IoT 허브
Microsoft
기본 사항 네트워크 관리 태그 검토 + 만들기
각 IoT 허브는 특정 계층에서 특정 장치 수로 프로비저닝됩니다. 계층과 장치 수에 따라 보낼 수 있는 메시지의 최대 일일 할당량이 결정됩니다. 자세한 정보
스케일링 계층 및 단위
가격 및 크기 계층 * ①
S1: 표준 계층
솔루션에 적합한 IoT 허브 계층을 선택하는 방법 알아보기
S1 IoT 허브 장치 수 ①
IoT 허브를 스케일링할 수 있는 방법을 결정합니다. 요구가 증가하는 경우 나중에 변경할 수 있습니다.
Defender for IoT
해제
Microsoft Defender for IoT는 Azure IoT Hub에 대한 위협 보호 계층을 추가하는 별도의 서비스입니다. IoT Edge 및 장치. 이 서비스는 별도로 청구됩니다. Defender for IoT는 IoT Hub와 다른 지리적 위치 내에서 데이터를 처리하고 저장할 수 있습니다. 자세히 알아보기
가격 및 크기 계층 ①
S1
디바이스-클라우드 메시지 ①
사용
일별 메시지 ①
400,000
메시지 라우팅 ①
사용
월별 비용
28116.25 KRW
클라우드 - 디바이스 명령 ①
사용
Defender for IoT ①
사용 안 함
IoT Edge ①
사용
디바이스 관리 ①
사용
역할 기반 액세스 제어
권한 모델을 Azure RBAC(역할 기반 액세스 제어)로만 변경하거나 공유 액세스 정책과 RBAC의 조합으로 변경합니다.
자세한 정보
RBAC만
공유 액세스 정책 + RBAC
인스턴스 내에서 요소를 관리하려면 사용자가 IoT Hub 데이터 API에 액세스해야 합니다. API에 대한 모든 권한을 부여하려면 아래에서 제안된 역할을 선택하세요. 액세스 제어(IAM)를 사용하여 나중에 적절한 역할을 선택할 수도 있습니다.
자세한 정보
나를 IoT Hub 데이터 기여자 역할에 할당 ①
고급 설정
스케일링
디바이스-클라우드 파티션 ①
4

7. 새로 만든 [IoT Hub] 블레이드로 이동한 후 [디바이스 관리 - 장치]를 클릭합니다. 앞서 추가한 Raspberry Pi 시뮬레이터를 디바이스로 추가하기 위해 [디바이스 추가]를 클릭합니다.

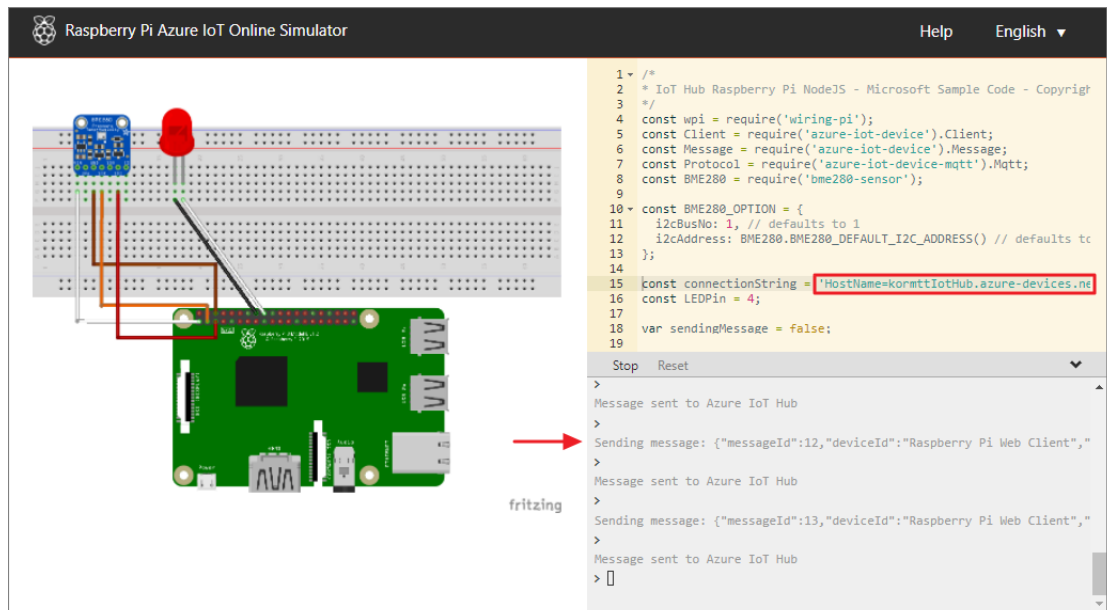
kormttlotHub | 장치
IoT Hub
검색(Ctrl+/)
IoT Hub에서 디바이스를 보고, 만들고, 삭제하고, 업데이트합니다.
디바이스 관리
장치
IoT Edge
구성
업데이트
쿼리
허브 설정
기본 제공 엔드포인트
디바이스 이름
디바이스 ID 입력
디바이스 찾기
쿼리를 사용하여 찾기
+ 디바이스 추가
새로 고침
삭제
장치 ID
상태
마지막 상태 ...
인증 ...
클라...
표시할 IoT 디바이스가 없습니다.

8. [디바이스 만들기] 블레이드에서 아래와 같이 구성하고 [저장]을 클릭합니다.
 - 디바이스 ID: RaspberryPi-0001
 - 다른 설정은 모두 기본값을 사용합니다.

9. [IoT Hub] 블레이드의 [디바이스 관리 - 장치]로 이동한 후 [새로 고침]을 클릭합니다. 새로 추가한 디바이스가 표시되는지 확인하고 이를 클릭합니다.

10. [RaspberryPi-0001] 블레이드에서 기본 연결 문자열 값을 클립보드로 복사합니다.

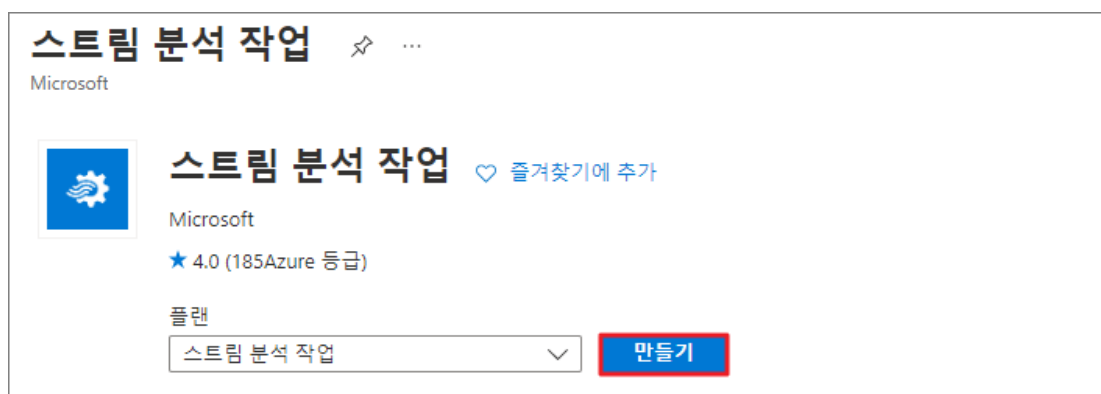
11. Raspberry Pi 시뮬레이터 탭으로 이동합니다. `const connectionString`에 위에서 복사한 값을 붙여 넣고 [Run]을 클릭합니다. 아래와 같이 메시지가 IoT Hub로 전송되는 것을 확인합니다.



TASK 02. 입력(IoT Hub)과 출력(Blob Storage) 구성

이제 IoT Hub가 설정되었고 디바이스에서 들어오는 스트림이 있는 상태입니다. 이제 Stream Analytics 작업을 만들어야 합니다.

1. Azure 포털에서 [리소스 만들기]를 클릭한 후 "스트림 분석 작업"을 검색한 후 클릭합니다. [스트림 분석 작업] 블레이드에서 [만들기]를 클릭합니다.



2. [새 Stream Analytics 작업] 블레이드에서 아래와 같이 구성한 후 [만들기]를 클릭합니다.
 - 작업 이름: processlotData
 - 리소스 그룹: 05_streamAnalyticsRg
 - 위치: Korea Central
 - 호스팅 환경: 클라우드
 - 스트리밍 단위: "1"로 설정합니다. 스트리밍 단위는 비용에 영향을 줍니다. Stream Analytics 비용은 스트리밍 단위 X 실행 시간 X 시간당 가격으로 책정됩니다.
 - 내 Storage 계정에서 이 작업에 필요한 모든 프라이빗 데이터 자산을 보호합니다: 선택하지 않습니다.

새 Stream Analytics 작업 ...

이렇게 하면 새 Stream Analytics 작업이 만들어집니다. Azure Stream Analytics 청구 모델에 따라 요금이 청구됩니다. 자세히 알아보세요. →

작업 이름 *
processIotData ✓

구독 *
Azure Pass - 스폰서쉽 ✓

리소스 그룹 *
05_streamAnalyticsRg ✓
[새로 만들기](#)

위치 *
Korea Central ✓

호스팅 환경 ①
클라우드 에지

스트리밍 단위(1~192) ①
○----- 1

☐ 내 Storage 계정에서 이 작업에 필요한 모든 프라이빗 데이터 자산을 보호합니다. ①

3. 새로 만든 [processIotData Stream Analytics] 작업 블레이드로 이동합니다. [개요]에서 작업의 시작 및 중지 메뉴가 있는 것을 확인합니다. Stream Analytics 작업은 생성되면 중지된 상태이며 이 상태에서는 비용이 발생하지 않습니다.

processIotData Stream Analytics 작업

검색(Ctrl+/) << **▶ 시작** ☐ 중지 **🗑 삭제** **🔄 새로 고침**

작업을 시작하려면 입력을 추가해야 합니다. →

기본 정보 [JSON 보기](#)

리소스 그룹 (이동)	만들어짐
05_streamAnalyticsRg	2022년 4월 4일 월요일 오후 10:20:54
상태	시작됨
Created	-
위치	출력 워터마크
Korea Central	-
구독 (이동)	클러스터
Azure Pass - 스폰서쉽	공유
구독 ID	호스팅 환경
1e38264e-a5cb-4cb2-b882-43f52ddb6f3	클라우드

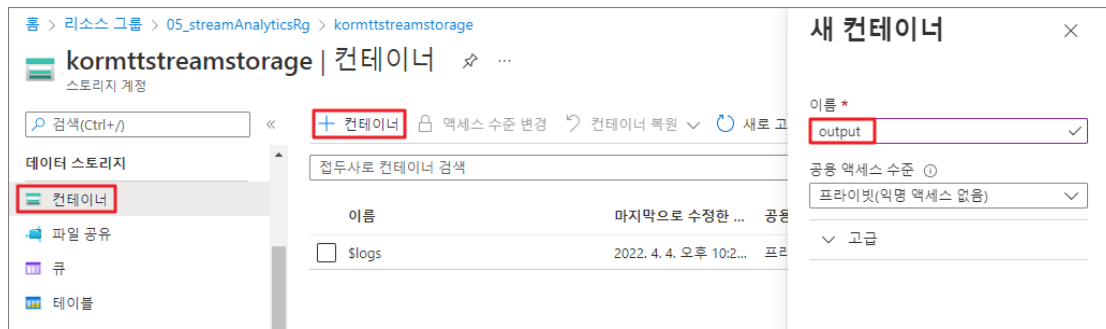
4. 이제 Stream Analytics의 출력을 설정해야 합니다. 이 실습에서는 Blob Storage를 Stream Analytics의 출력으로 사용합니다.
5. Azure 포털에서 [리소스 만들기]를 클릭한 후 "스토리지 계정"을 검색하고 클릭합니다. [스토리지 계정] 블레이드에서 [만들기]를 클릭합니다.



6. [저장소 계정 만들기] 블레이드의 [기본] 탭에서 아래와 같이 구성하고 [검토 + 만들기]를 클릭합니다. [검토 + 만들기] 탭에서 [만들기]를 클릭합니다.

- [프로젝트 정보 - 리소스 그룹]: 05_streamAnalyticsRg
- [인스턴스 정보 - 스토리지 계정 이름]: 중복되지 않는 고유한 이름을 입력합니다.
- [인스턴스 정보 - 지역]: (Asia Pacific) Korea Central
- [인스턴스 정보 - 성능]: 표준
- [인스턴스 정보 - 중복]: LRS(로컬 중복 스토리지)

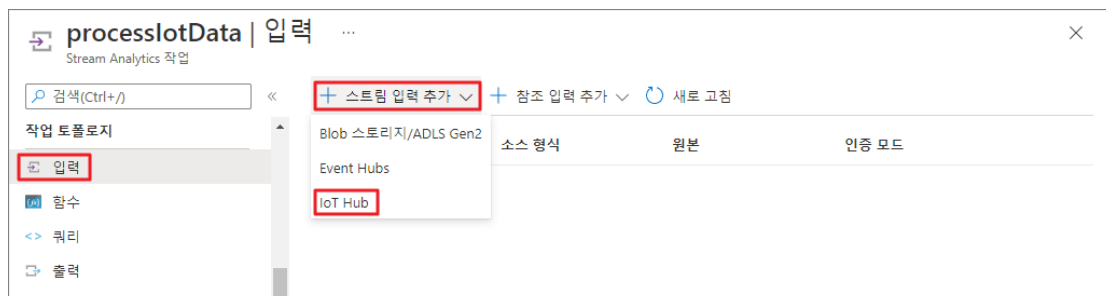
7. 새로 만든 [스토리지 계정] 블레이드로 이동합니다. [데이터 스토리지 - 컨테이너]에서 [컨테이너]를 클릭합니다. [새 컨테이너]에서 이름에 "output"을 입력하고 [만들기]를 클릭합니다.



TAKS 03. Stream Analytics 작업 구성

이 작업에서는 앞서 만들었던 IoT Hub를 입력으로 사용할 수 있도록 Stream Analytics에서 입력을 정의합니다.

1. [processIotData Stream Analytics 작업] 블레이드의 [작업 토폴로지 - 입력]으로 이동한 후 메뉴에서 [스트림 입력 추가 - IoT Hub]를 클릭합니다.



2. [IoT Hub]에서 아래와 같이 구성한 후 [저장]을 클릭합니다.

- 입력 별칭: iotInput
- IoT Hub: 앞서 만들었던 IoT Hub를 선택합니다.
- 다른 모든 옵션은 기본값으로 설정합니다.

IoT Hub
새 입력

입력 별칭 *
iotInput ✓

☐ IoT Hub 설정을 수동으로 제공
☒ 구독에서 IoT Hub 선택

구독
Azure Pass - 스폰서쉽

IoT Hub * ⓘ
kormttIoTHub

소비자 그룹 * ⓘ
\$Default

공유 액세스 정책 이름 * ⓘ
iothubowner

공유 액세스 정책 키 ⓘ
.....

엔드포인트 ⓘ
메시지

파티션 키 ⓘ

이벤트 serialization 형식 * ⓘ
JSON

인코딩 ⓘ
UTF-8

이벤트 압축 유형 ⓘ
없음

3. 새 스트림 입력을 추가하면 추가한 입력 스트림에 대한 연결 테스트가 진행되고 연결 테스트가 성공한 것을 확인할 수 있습니다. 연결 테스트가 다시 필요한 경우 입력 스트림의 [테스트] 아이콘을 클릭하여 연결을 다시 테스트할 수 있습니다.



TASK 04. Blob Storage로 출력 구성

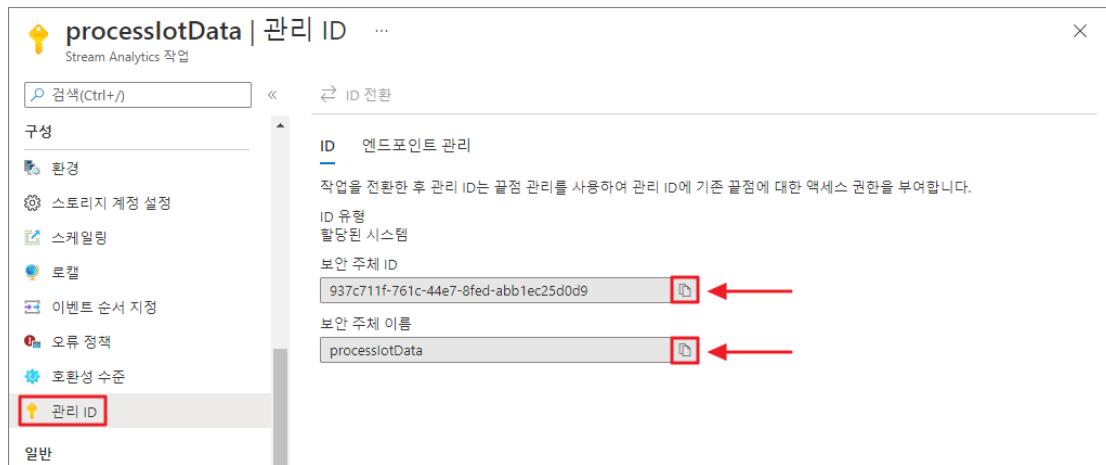
이제 Stream Analytics의 출력으로 Blob Storage를 구성해야 합니다. Blob Storage로 출력하기 위해 **Managed ID**를 구성하게 되면 Stream Analytics 작업에 연결 문자열을 사용하는 대신 스토리지 계정에 대한 직접 액세스 권한을 부여합니다. 또한 향상된 보안 외에도 가상 네트워크 내에 있는 스토리지 계정에 액세스할 수도 있습니다.

1. [processIotData Stream Analytics 작업] 블레이드의 [구성 - 관리 ID]로 이동합니다. [ID 선택]을 클릭한

후 [관리 ID]에서 "시스템 할당 ID"를 선택하고 [저장]을 클릭합니다.



2. [구성 - 관리 ID]에서 표시되는 보안 주체 ID와 보안 주체 이름 값을 메모장에 복사합니다. 이 ID가 스토리지 계정에 추가하게 되는 ID와 동일하지 추후 확인해야 합니다.



3. 위에서 추가한 보안 주체 ID는 스토리지 계정 혹은 컨테이너 수준에 대해 액세스를 설정할 수 있습니다. 이 작업에서는 스토리지 계정에 대한 액세스 권한을 부여합니다.
4. [스토리지 계정] 블레이드의 [액세스 제어(IAM)]로 이동합니다. 메뉴에서 [추가 - 역할 할당 추가]를 클릭합니다.



5. [역할 할당 추가] 블레이드의 [역할] 탭에서 "Storage Blob"을 검색한 후 "Storage Blob 데이터

Contributor"를 선택하고 [다음]을 클릭합니다.

역할 할당 추가

역할 정의는 권한 컬렉션입니다. 기본 제공 역할을 사용하거나 사용자 지정 역할을 만들 수 있습니다. 자세한 정보 > 클래식 환경 사용 ①

storage blob

형식: 모두 범주: 모두

40개 역할 중 4개 표시 중

이름 ↑↓	설명 ↑↓	형식 ↑↓	범주 ↑↓	세부 정보
Storage Blob 데이터 Contributor	Azure Storage Blob 컨테이너 및 데이터에 대한 ...	BuiltInRole	스토리지	보기
Storage Blob 데이터 Reader	Azure Storage Blob 컨테이너 및 데이터에 대한 ...	BuiltInRole	스토리지	보기
Storage Blob 데이터 소유자	POSIX 액세스 제어 할당을 포함하여 Azure Stora...	BuiltInRole	스토리지	보기
Storage Blob 위임자	SAS 토큰에 서명하는 데 사용할 수 있는 사용자 ...	BuiltInRole	스토리지	보기

6. [구성원] 탭에서 아래와 같이 구성한 후 [검토 + 할당]을 클릭합니다. [검토 + 할당] 탭에서 [검토 + 할당]을 클릭합니다.

- 다음에 대한 액세스 할당: 사용자, 그룹 또는 서비스 주체
- [구성원 선택]을 클릭합니다. [구성원 선택]에서 앞서 메모장에 복사한 Stream Analytics의 보안 주체 이름(processIotData)으로 검색한 후 이를 선택하고 [선택]을 클릭합니다.
- 추가된 구성원의 개체 ID 값이 앞서 메모장에 복사한 Stream Analytics의 보안 주체 ID와 일치하는지 확인합니다.

역할 할당 추가

선택한 역할: Storage Blob 데이터 Contributor

다음에 대한 액세스 할당: ☒ 사용자, 그룹 또는 서비스 주체 ☐ 관리 ID

구성원: + 구성원 선택

이름	개체 ID	유형
processIotData	937c711f-761c-44e7-8fed-abb1ec25d0d9	앱

Description: 선택 사항

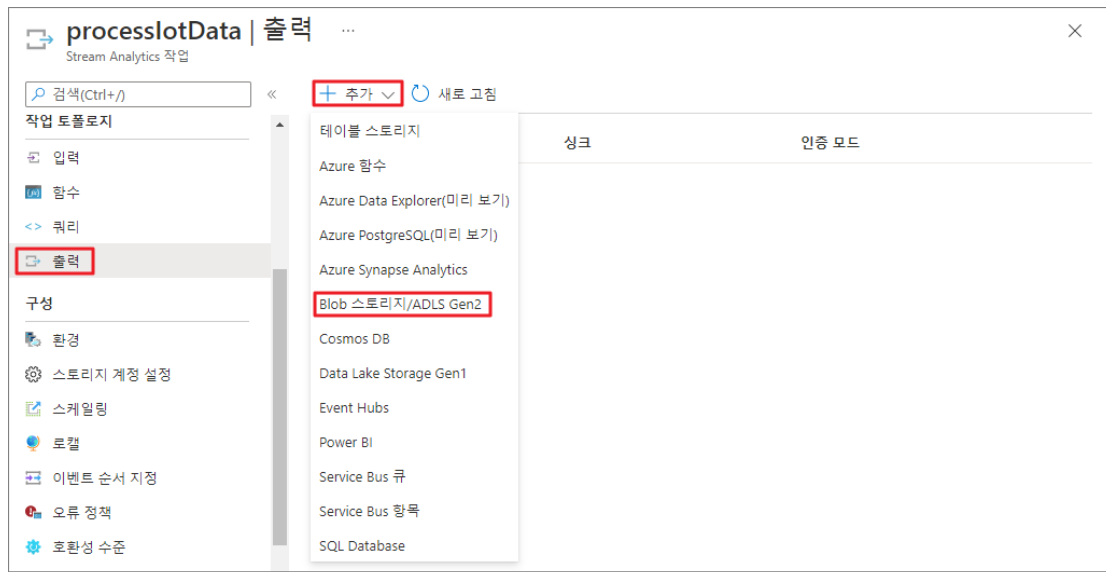
구성원 선택

선택 ①: processIotData

사용자, 그룹 또는 서비스 보안 주체가 없습니다.

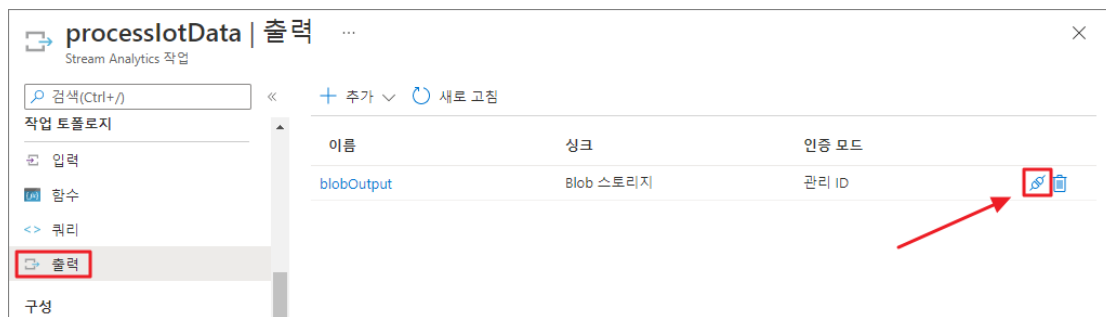
선택한 구성원: processIotData 제거

7. [processIotData Stream Analytics 작업] 블레이드의 [작업 토폴로지 - 출력]으로 이동합니다. 메뉴에서 [추가 - Blob 스토리지/ADLS Gen2]를 클릭합니다.



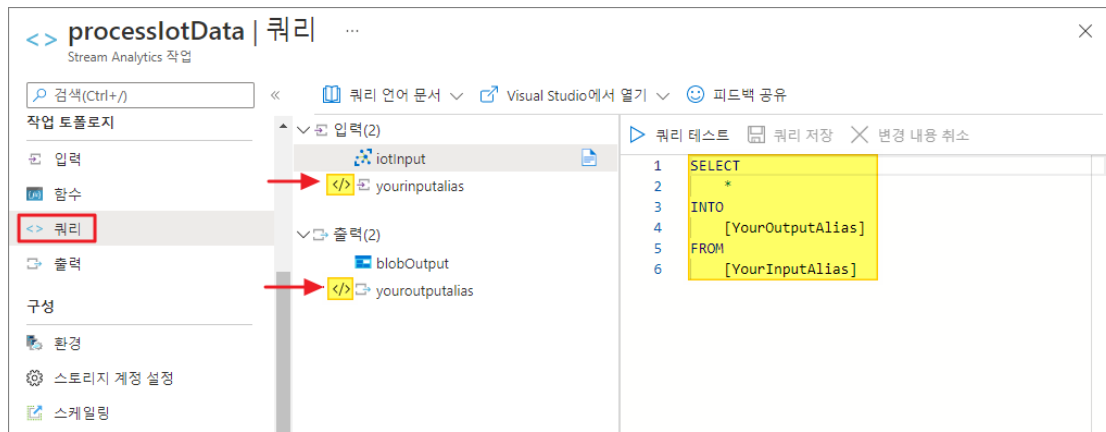
8. [Blob 스토리지/ADLS Gen2]에서 아래와 같이 구성한 후 [저장]을 클릭합니다.
- 출력 별칭: blobOutput
 - 구독에서 Blob 스토리지/ADLS Gen2 선택
 - 스토리지 계정: 앞서 만들었던 스토리지 계정을 선택합니다.
 - 컨테이너: "기존 항목 사용"을 선택하고 "output" 컨테이너를 선택합니다.
 - 인증 모드: 관리 ID: 할당된 시스템
 - 경로 패턴: {date}를 입력합니다. {datetime}, {date}/{time}과 같이 다른 경로 패턴을 지정할 수도 있습니다.
 - 날짜 형식: YYYY/MM/DD
 - 이벤트 serialization 형식: JSON. 이 외에도 CSV, Avro, Parquet 등의 형식을 선택할 수 있습니다.
 - 다른 옵션은 모두 기본값을 선택합니다.

9. [작업 토폴로지 - 출력]에서 새로 추가한 출력이 표시되는지 확인합니다. [테스트] 아이콘을 클릭하여 Blob Storage로 연결 테스트가 성공하는지 확인합니다.



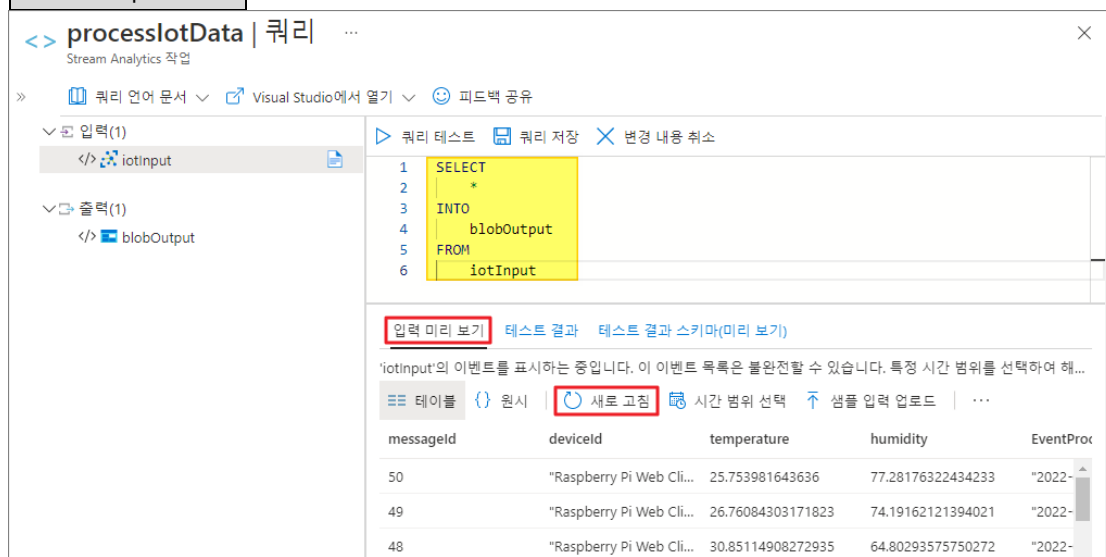
10. [processIotData Stream Analytics 작업] 블레이드의 [작업 토폴로지 - 쿼리]로 이동합니다. 다음과 같은 내용을 확인할 수 있습니다.

- 여기에서 개발과 관련하여 원하는 모든 작업(사용자 지정 함수 포함)을 구성할 수 있습니다.
- 쿼리 창에 표준 SQL 문이 표시되고 왼쪽 패널에 있는 입력과 출력을 사용하고 있다는 것을 확인할 수 있습니다.
- 왼쪽 패널에서 현재 사용 중인 입력과 출력에는 사용 중임을 표시하는 `</>` 아이콘이 표시됩니다.

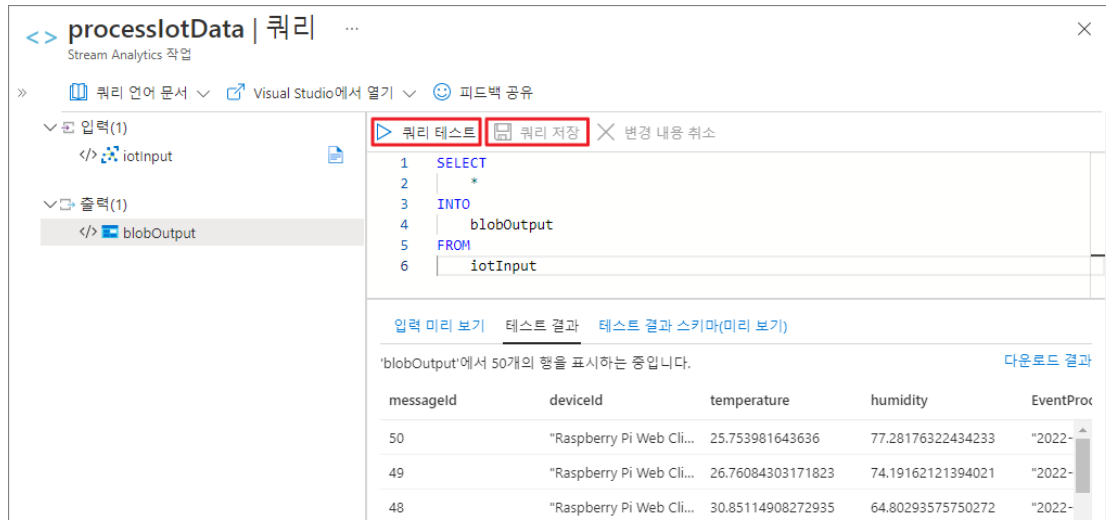


11. 쿼리 창에 다음과 같은 쿼리를 입력하여 **iotInput** 입력의 내용 모두를 **blobOutput** 출력으로 복사하도록 작성합니다. 쿼리를 테스트하기 전에 페이지 하단의 [입력 미리 보기] 탭에서 [새로 고침]을 클릭하여 입력 데이터의 샘플 데이터가 어떤 것인지 검토합니다.

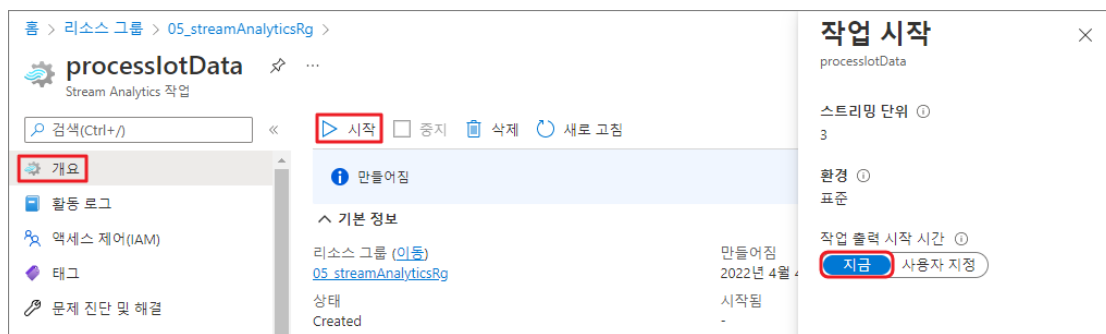
```
SELECT
*
INTO
blobOutput
FROM
iotInput
```



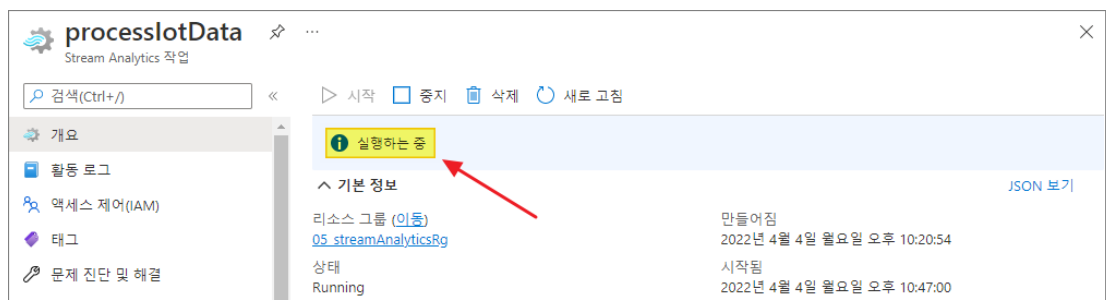
12. [쿼리 테스트]를 클릭하여 출력으로 데이터가 성공적으로 복사되는지 확인합니다. **SELECT *** 를 사용했기 때문에 IoT Hub에서 모든 데이터를 가져온 다음 Blob Storage로 스트리밍하는 것을 확인할 수 있습니다. [쿼리 저장]을 클릭합니다.



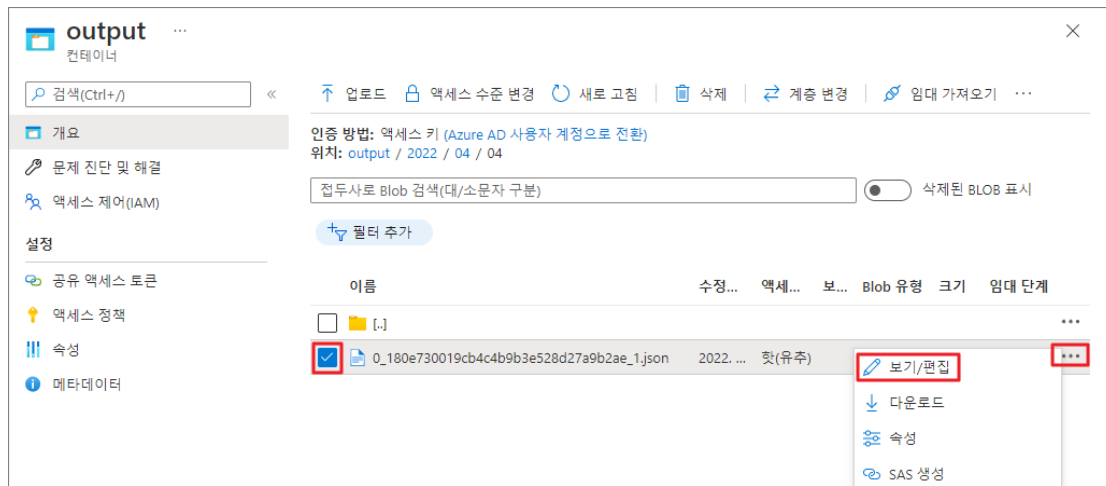
13. [processIotData Stream Analytics 작업] 블레이드의 [개요]로 이동한 후 [시작]을 클릭합니다. [작업 시작]에서 [지금]을 선택한 후 [시작]을 클릭합니다.



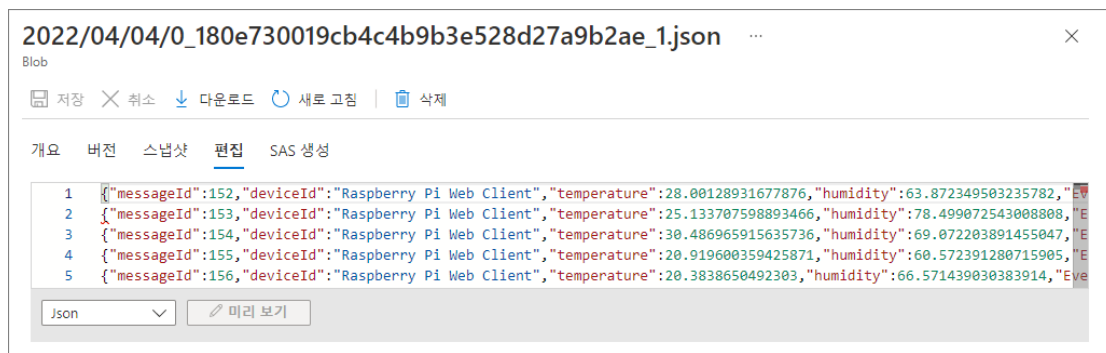
14. [processIotData Stream Analytics 작업] 블레이드의 [개요]에서 [새로 고침]을 클릭하고 "실행하는 중" 메시지가 표시되는지 확인합니다.



15. [스토리지 계정] 블레이드의 [데이터 스토리지 - 컨테이너]로 이동합니다. output 컨테이너를 클릭하고 연도, 월, 일 순서의 가상 디렉터리로 이동합니다. 생성된 JSON 파일을 선택한 후 [... - 보기/편집]을 클릭합니다.



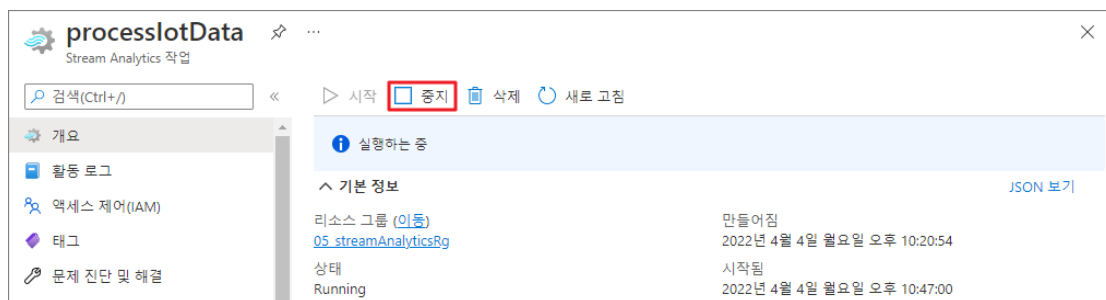
16. JSON 파일 페이지에서 [편집] 탭으로 이동한 후 스트리밍 데이터가 표시되는지 확인합니다. 페이지를 새로 고칠 때마다 계속 데이터가 추가되는 것 또한 확인합니다.



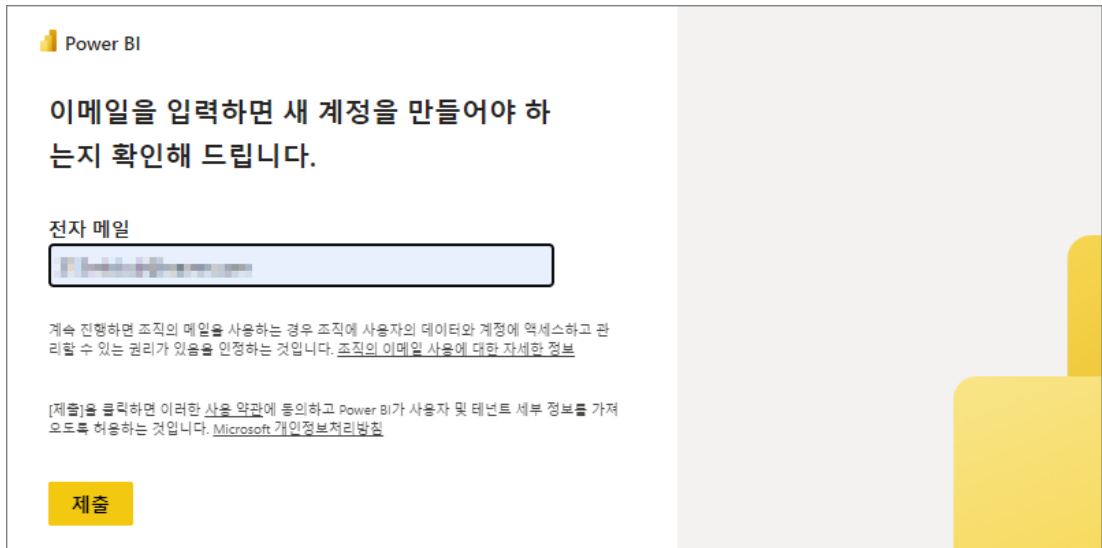
TASK 05. Power BI로 집계 출력 구성

이 작업에서는 유사한 SQL 쿼리를 사용하여 IoT Hub의 데이터를 Power BI로 출력합니다. 따라서 Power BI를 새 출력으로 추가하고 특정 시간 간격으로 입력된 데이터의 평균 값을 계산한 다음 이 값을 Power BI로 스트리밍합니다.

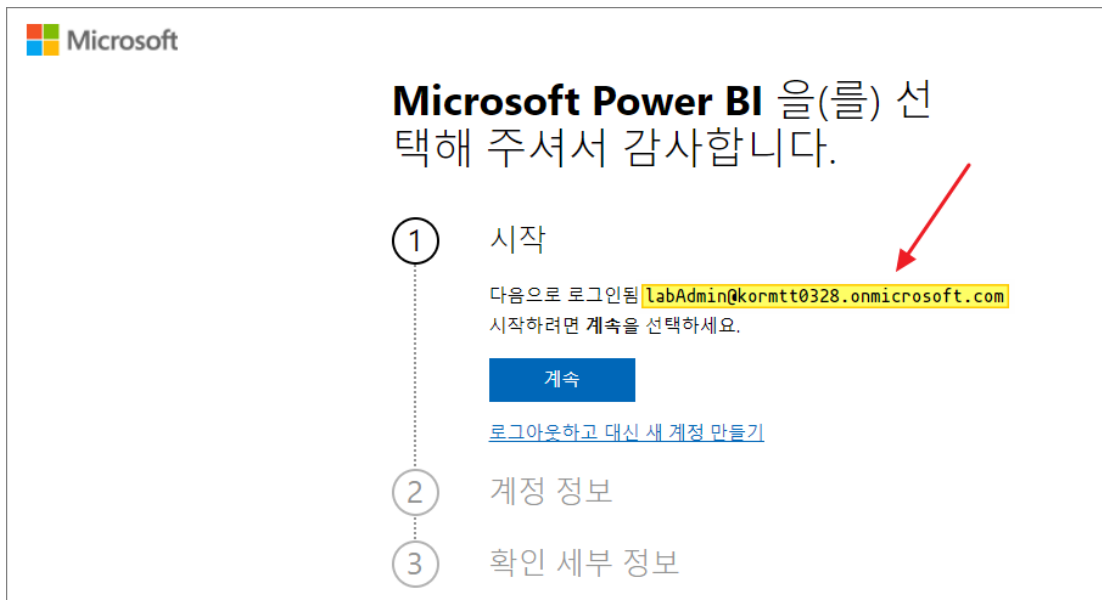
1. [processIotData] Stream Analytics 작업 블레이드의 [개요]에서 [중지]를 클릭합니다. [Stream Analytics 작업 중지]에서 [예]를 클릭합니다. Stream Analytics 작업을 중지하지 않으면 새 입력이나 출력을 추가할 수 없습니다.



2. 브라우저에서 새 탭을 열고 <https://powerbi.com/> 사이트로 이동합니다. 임의의 전자 메일 계정을 입력하고 [제출]을 클릭합니다.



3. [Microsoft Power BI 을(를) 선택해 주셔서 감사합니다 - 시작] 페이지에서 자신의 Azure 구독 계정으로 로그인되어 있는지 확인한 후 [계속]을 클릭합니다.



4. [계정 정보] 페이지에서 로그인 정보를 확인하고 전화 번호를 입력한 후 [시작]을 클릭합니다.

Microsoft Power BI 을(를) 선택해 주셔서 감사합니다.

- 계정 설정
- 계정 정보

안녕하세요, **lab** 님.
시작하기 전에 추가 정보가 필요합니다

labAdmin@kormtt0328.onmicrosoft.com

국가 또는 지역
대한민국

전화 번호

Microsoft에서 내 평가판에 대해 문의할 수 있음을 알고 있습니다.

☐ 기업 및 조직을 위한 솔루션 및 여타 Microsoft 제품과 서비스에 대한 개인 설정된 정보, 팁, 혜택을 받아보고 싶습니다. [개인정보처리방침](#).

☐ 제품 및 서비스에 대한 관련 정보를 받을 수 있도록 Microsoft에서 내 정보를 특정 파트너와 공유하고 싶습니다. 자세한 내용은 [개인정보처리방침](#)을 참조하세요.

시작을 선택하면, [사용 약관](#) 및 [개인정보처리방침](#)에 동의하게 됩니다.

시작

5. [확인 세부 정보] 페이지에서 [시작]을 클릭합니다.

Microsoft Power BI 을(를) 선택해 주셔서 감사합니다.

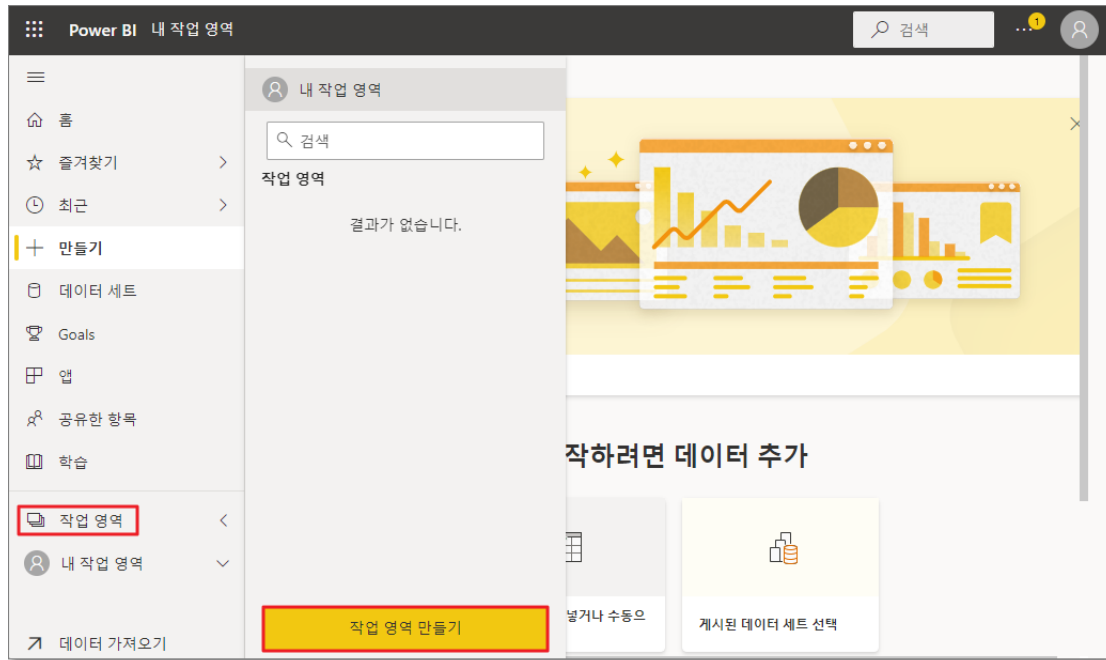
- 계정 설정
- 님, 계정을 만들어 주셔서 감사합니다.
- 확인 세부 정보

가입해 주셔서 감사합니다.

사용자 ID는 **labAdmin@kormtt0328.onmicrosoft.com**입니다.

시작

6. 왼쪽 네비게이션 메뉴에서 [작업 영역]을 클릭한 후 [작업 영역 만들기]를 클릭합니다. [Power BI Pro로 업그레이드] 창이 표시되면 [무료 체험]을 클릭합니다. [마지막 단계 하나만] 창이 표시되면 정보를 입력하고 필수 사항에 동의한 후 [평가판 시작]을 클릭합니다.



7. [작업 영역 만들기]에서 작업 영역 이름에 "streamAnalytics"를 입력하고 [저장]을 클릭합니다.

작업 영역 만들기

작업 영역 이미지

업로드
삭제

작업 영역 이름

streamAnalytics

사용 가능

설명

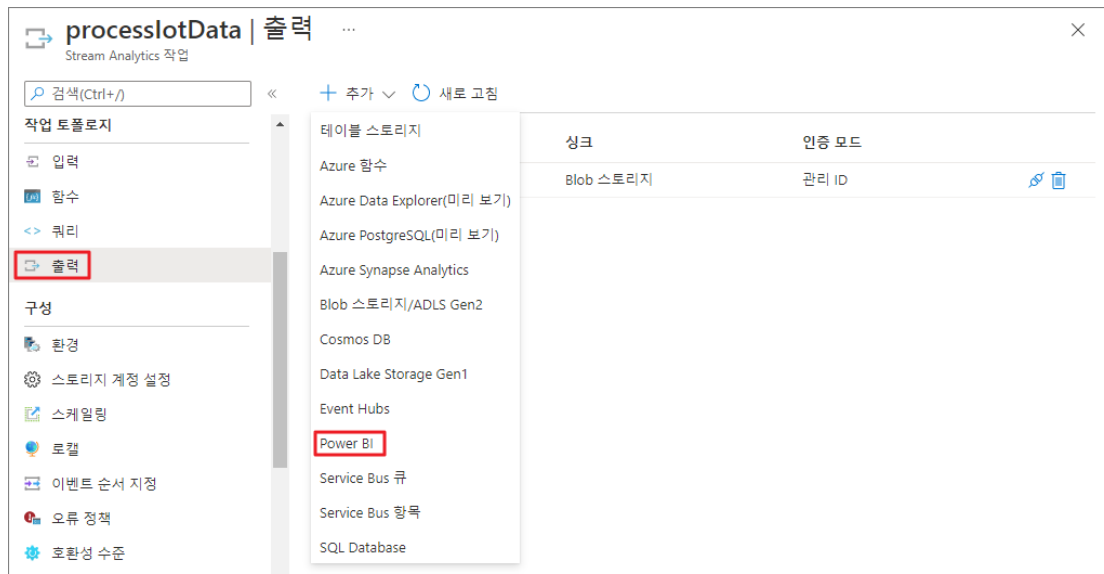
이 작업 영역에 대해 설명합니다.

[작업 영역 설정에 대한 자세한 정보](#)

고급 ▾

저장 취소

8. [processIotData Stream Analytics 작업] 블레이드의 [작업 토폴로지 - 출력]으로 이동합니다. 메뉴에서 [추가 - Power BI]를 클릭합니다.



9. [Power BI]에서 아래와 같이 구성한 후 [저장]을 클릭합니다.

- 출력 별칭: powerBiOutput
- 구독에서 그룹 작업 영역 선택
- 그룹 작업 영역: 앞서 만들었던 Power BI Workspace를 선택합니다.
- 인증 모드: "관리 ID: 할당된 시스템". "사용자 토큰"을 선택하는 경우 별도의 권한 부여 작업을 진행해야 합니다.
- 데이터 세트 이름: iotData
- 테이블 이름: Temperature

Power BI
새 출력

출력 별칭 *

powerBiOutput ✓

☐ 그룹 작업 영역 설정을 수동으로 제공
☒ 구독에서 그룹 작업 영역 선택

그룹 작업 영역 *

streamAnalytics ▼

인증 모드

관리 ID: 할당된 시스템 ▼

저장을 클릭하면 이 Stream Analytics 작업의 관리 ID에 기여자 역할이 부여됩니다. 권한 부여에 실패하면 [여기](#)의 수동 권한 부여 단계에 따라 수행하세요.

데이터 세트 이름 * ①

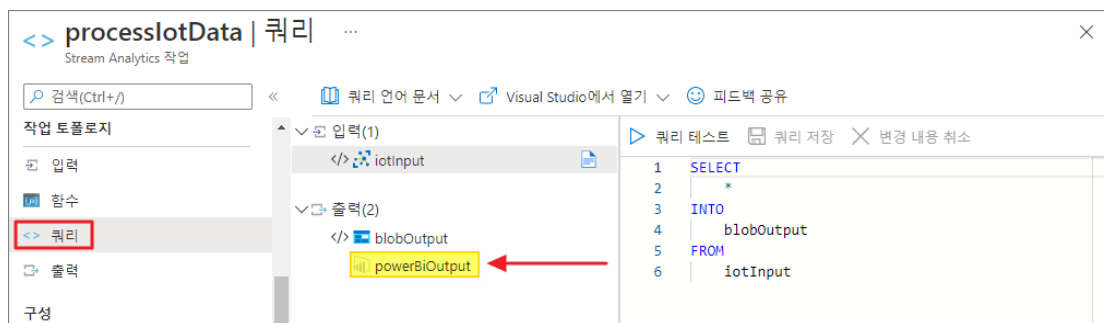
iotData ✓

테이블 이름 *

Temperature ✓

i 관리 ID 액세스는 Power BI Pro 구독 및 업그레이드된 구독 또는 V2 작업 영역에서만 작동합니다. 이 기능에 대한 자세한 내용은 [설명서](#)를 참조하세요.

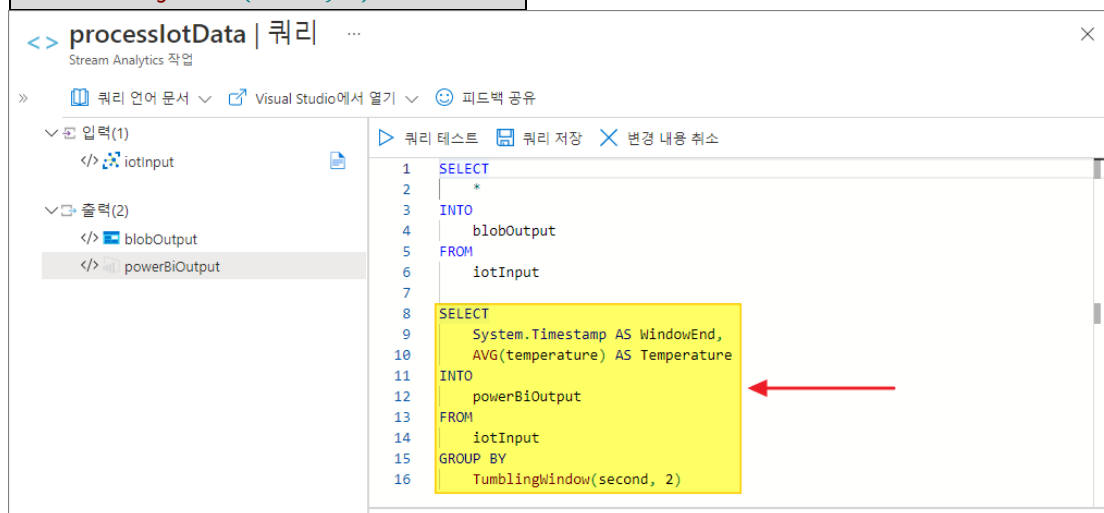
10. Power BI 연결이 구성되고 연결 테스트가 성공적으로 완료되었는지 확인합니다. [processIotData Stream Analytics 작업] 블레이드의 [작업 토폴로지 - 쿼리]로 이동합니다. 출력으로 추가한 Power BI 출력이 표시되는 것을 확인합니다.



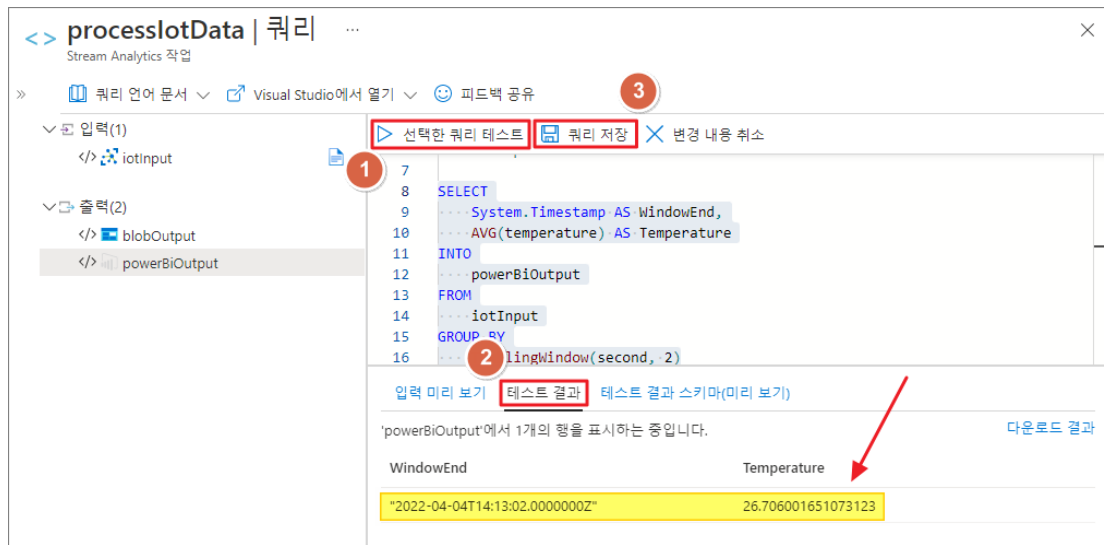
11. 쿼리 창에서 다음과 같은 쿼리를 추가합니다.

- 쿼리 창에서 여러 쿼리를 사용할 수 있으며 여러 쿼리는 여러 입력과 출력을 사용할 수 있습니다. 즉 이전에 작성한 쿼리와 새로 작성한 쿼리는 함께 작동하게 됩니다.
- SELECT: 평균값 계산이 완료된 시간을 확인하기 위해 `System.Timestamp` 값을 가져와 `WindowEnd` 열 이름으로 표시하고, `temperature`의 평균(AVG) 값을 `Temperature` 이름으로 표시합니다.
- INTO: SELECT 문에서 표시되는 값을 앞서 지정한 `powerBiOutput`으로 출력합니다.
- FROM: 입력값으로 `iotInput` 입력을 사용합니다.
- GROUP BY: 평균값을 계산하기 위해 `window` 기반 함수를 사용할 것이기 때문에 `GROUP BY`가 필요합니다. `TumblingWindow` 함수를 사용하여 매 2초 간격으로 평균값을 구합니다.

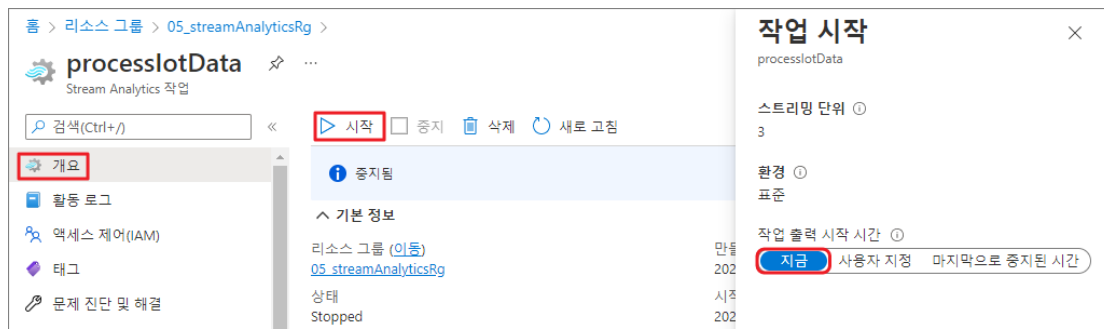
```
SELECT
    System.Timestamp AS WindowEnd,
    AVG(temperature) AS Temperature
INTO
    powerBiOutput
FROM
    iotInput
GROUP BY
    TumblingWindow(second, 2)
```



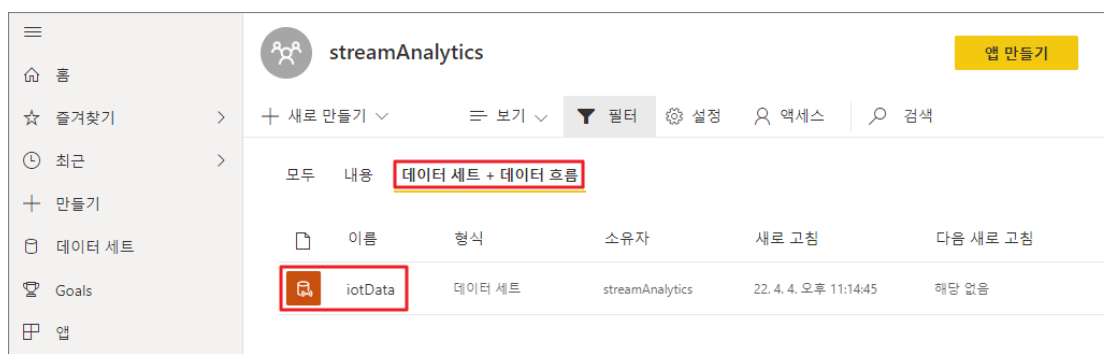
12. 쿼리창에서 새로 추가한 쿼리를 선택한 후 메뉴에서 [선택한 쿼리 테스트]를 클릭합니다. 평균 값과 평균 값이 계산된 시스템 시간이 출력되는 것을 확인합니다. 쿼리 결과가 정상적으로 출력되면 [쿼리 저장]을 클릭합니다.



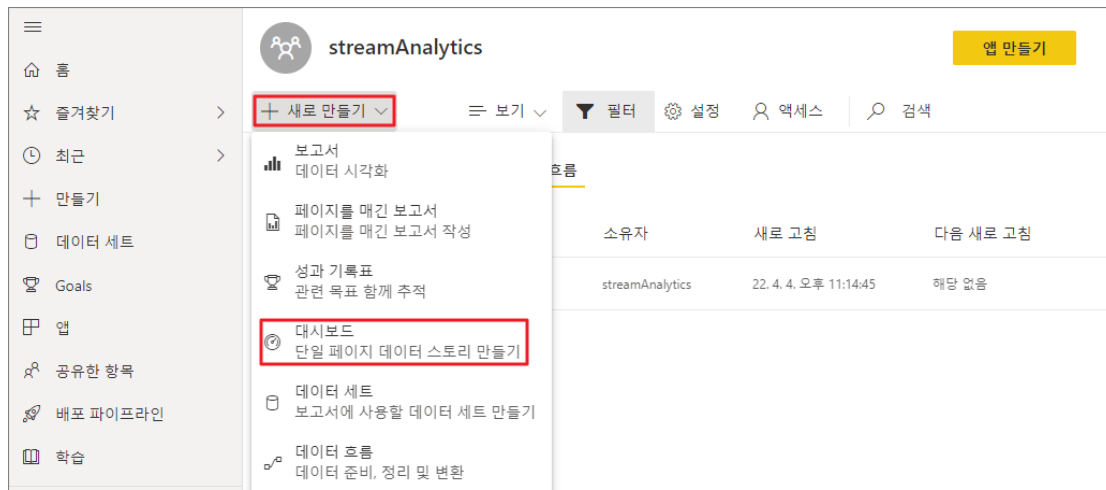
13. [processIotData Stream Analytics 작업] 블레이드의 [개요]로 이동한 후 [시작]을 클릭합니다. [작업 시작]에서 [지금]을 선택하고 [시작]을 클릭합니다.



14. Stream Analytics 작업이 시작되면 Power BI 포털로 이동합니다. Power BI의 왼쪽 네비게이션 메뉴에서 [작업 영역 - streamAnalytics] 작업 영역을 선택합니다. [데이터 세트 + 데이터 흐름] 탭에서 아래와 같이 Stream Analytics에서 지정했던 데이터 세트 이름이 표시되는 것을 확인할 수 있습니다.



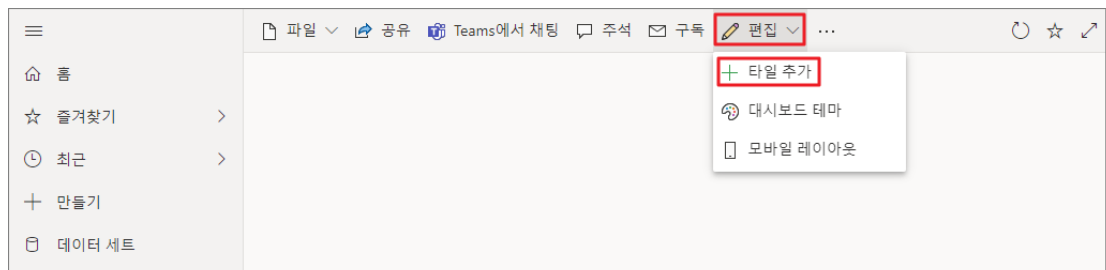
15. Power BI에서 [새로 만들기 - 대시보드]를 클릭합니다.



16. [대시보드 만들기] 창에서 이름에 "IoT Temperature"를 입력하고 [Create]를 클릭합니다.



17. 대시보드의 메뉴에서 [편집 - 타일 추가]를 클릭합니다.



18. [타일 추가 - 원본 선택]에서 [사용자 지정 스트리밍 데이터] 타일을 선택하고 [다음]을 클릭합니다.

타일 추가

원본 선택

미디어

웹 콘텐츠

이미지

텍스트 상자

비디오

실시간 데이터

사용자 지정 스트리밍 데이터

다음

취소

19. [사용자 지정 스트리밍 데이터 타일 추가 - 스트리밍 데이터 세트 선택]에서 "iotData"를 선택하고 [다음]을 클릭합니다.

사용자 지정 스트리밍 데이터 타일 추가

스트리밍 데이터 세트 선택

+ 스트리밍 데이터 세트 추가

데이터 세트

데이터 세트 관리

뒤로

다음

취소

20. [사용자 지정 스트리밍 데이터 타일 추가 - 시각화 디자인]에서 아래와 같이 선택하고 [다음]을 클릭합니다.

- 시각화 형식: 카드
- 필드: [값 추가]를 클릭한 후 Temperature를 선택합니다.

사용자 지정 스트리밍 데이터 타일 추가

스트리밍 데이터 세트 선택 > 시각화 디자인

시각화 형식

카드

필드

Temperature

+ 값 추가

데이터 세트 관리

뒤로 다음 취소

21. [타일 세부 정보]에서 적절한 제목을 입력하고 [적용]을 클릭합니다.

타일 세부 정보

* 필수

자세히

☒ 제목 및 부제목 표시

제목

평균 온도(2초 간격)

부제목

기능

☐ 사용자 지정 링크 설정

링크 유형

☒ 외부 링크

☐ 현재 작업 영역의 대시보드 또는 보고서에 대한 링크

URL *

동일한 탭에서 사용자 지정 링크 열기

☐ 예

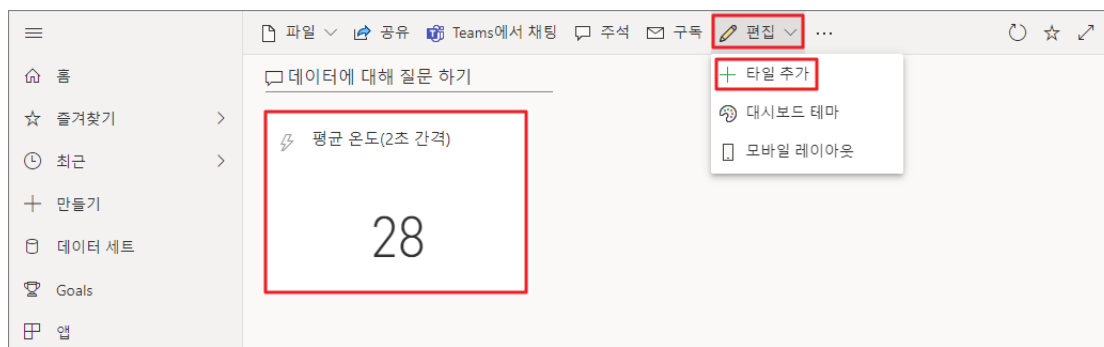
☒ 아니요

기본값 복원

기술 세부 정보

뒤로
적용
취소

22. 아래와 같이 추가한 카드 형식의 평균 온도 값이 표시되는 것을 확인합니다. [편집 - 타일 추가]를 클릭합니다.



23. [타일 추가 - 원본 선택]에서 [사용자 지정 스트리밍 데이터] 타일을 선택하고 [다음]을 클릭합니다.

타일 추가

원본 선택

미디어

 웹 콘텐츠

 이미지

 텍스트 상자

 비디오

실시간 데이터

 사용자 지정 스트리밍 데이터

다음
취소

24. [사용자 지정 스트리밍 데이터 타일 추가 - 스트리밍 데이터 세트 선택]에서 "iotData"를 선택하고 [다음]을 클릭합니다.

사용자 지정 스트리밍 데이터 타일 추가

스트리밍 데이터 세트 선택

+ 스트리밍 데이터 세트 추가

데이터 세트

데이터 세트 관리

뒤로
다음
취소

25. [사용자 지정 스트리밍 데이터 타일 추가 - 시각화 디자인]에서 아래와 같이 구성한 후 [다음]을 클릭합니다.

- 시각화 형식: 꺾은선형 차트
- 축: [값 추가]를 클릭한 후 WindowEnd를 선택합니다.
- 값: [값 추가]를 클릭한 후 Temperature를 선택합니다.
- 표시할 시간 창: 마지막 30분

사용자 지정 스트리밍 데이터 타일 추가

스트리밍 데이터 세트 선택 > 시각화 디자인

시각화 형식

꺾은선형 차트

축

WindowEnd

+ 값 추가

범례

+ 값 추가

값

Temperature

+ 값 추가

표시할 시간 창

마지막 30 분

데이터 세트 관리

뒤로 다음 취소

26. [타일 세부 정보]에서 적절한 제목을 입력하고 [적용]를 클릭합니다.

타일 세부 정보

* 필수

자세히

☒ 제목 및 부제목 표시

제목

평균 온도 기록

부제목

기능

☐ 사용자 지정 링크 설정

링크 유형

☒ 외부 링크

☐ 현재 작업 영역의 대시보드 또는 보고서에 대한 링크

URL *

동일한 탭에서 사용자 지정 링크 열기

☐ 예

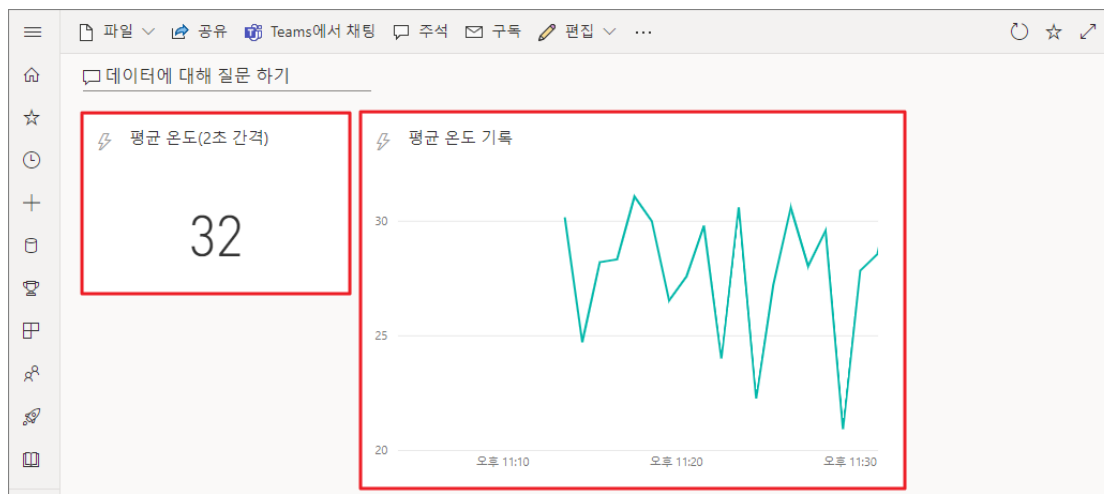
☒ 아니요

[기본값 복원](#)

[기술 세부 정보](#)

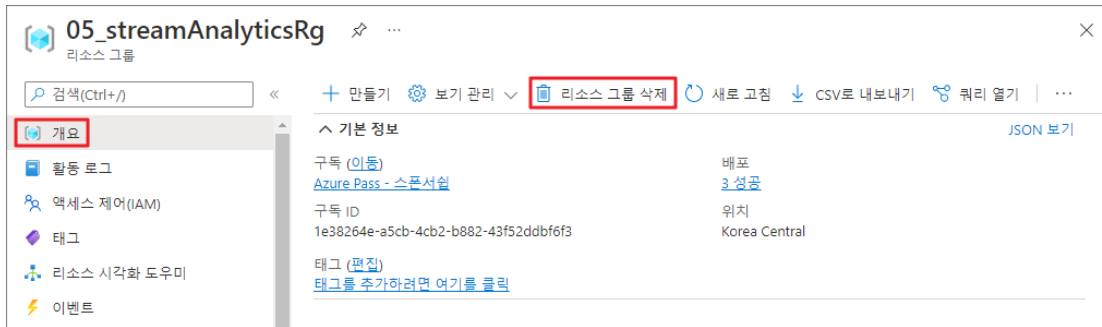
뒤로
적용
취소

27. 아래와 같이 IoT Hub를 통해 입력된 데이터의 실시간 평균값과 선형 차트 값을 확인할 수 있습니다.



TASK 06. 리소스 정리

1. Azure 포털에서 [05_streamAnalyticsRg 리소스 그룹] 블레이드로 이동한 후 메뉴에서 [리소스 그룹 삭제]를 클릭합니다.



2. 리소스 그룹 삭제 확인 창에서 리소스 그룹 이름을 입력한 후 [삭제]를 클릭합니다.

