

Azure Synapse Analytics for Gaming

Program Overview


Advanced Azure Data Analytics Day

게임사를 위한 Advanced Azure Data Analytics Training

Advanced Azure Data Analytics Day


매월 **선착순 10명**에게만 제공되는
심화 데이터 트레이닝, **AAA DAY**를 신청하세요!

신청하기




Azure Machine Learning

✓



Azure
Synapse
Analytics



Azure
Cognitive
Services

Advanced Azure Data Analytics Day

2월 24일(수)

Azure Synapse Analytics를 활용한 User Segmentation 기반의 Target Marketing

- 13:30 - 18:00
 - Azure Synapse Analytics 소개
 - Hands-on / Modeling Session

4월 28일(수)

Azure Machine Learning을 활용한 Abuser detection 진행

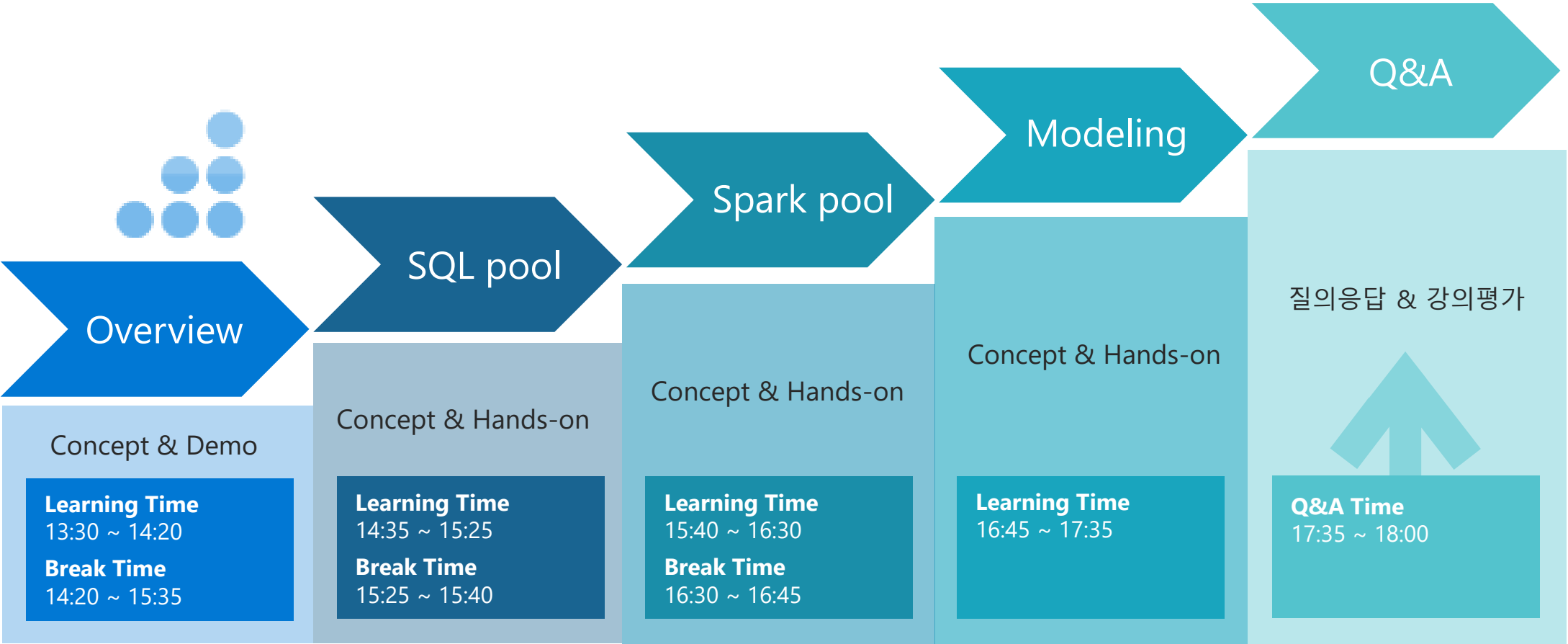
- 13:30 - 18:00
 - Azure Machine Learning 소개
 - Hands-on / Modeling Session

6월 30일(수)

Azure Cognitive Service를 활용한 Chatbot 및 in-game translation 개발

- 13:30 - 18:00
 - Azure Cognitive Service 소개
 - Hands-on / Modeling Session

Advanced Azure Data Analytics Day: Progress



Advanced Azure Data Analytics Day: Speakers



Jisun Kim

Microsoft
Data & AI



Hansol Song

Cloocus
DA



Jongsung Sim

Cloocus
DP



Yoonjeong Park

Cloocus
DS

Overview :
Azure Synapse Analytics
Concept & Demo



SQL pool :
Concept & Hands -on



Apache Spark pool :
Concept & Hands -on

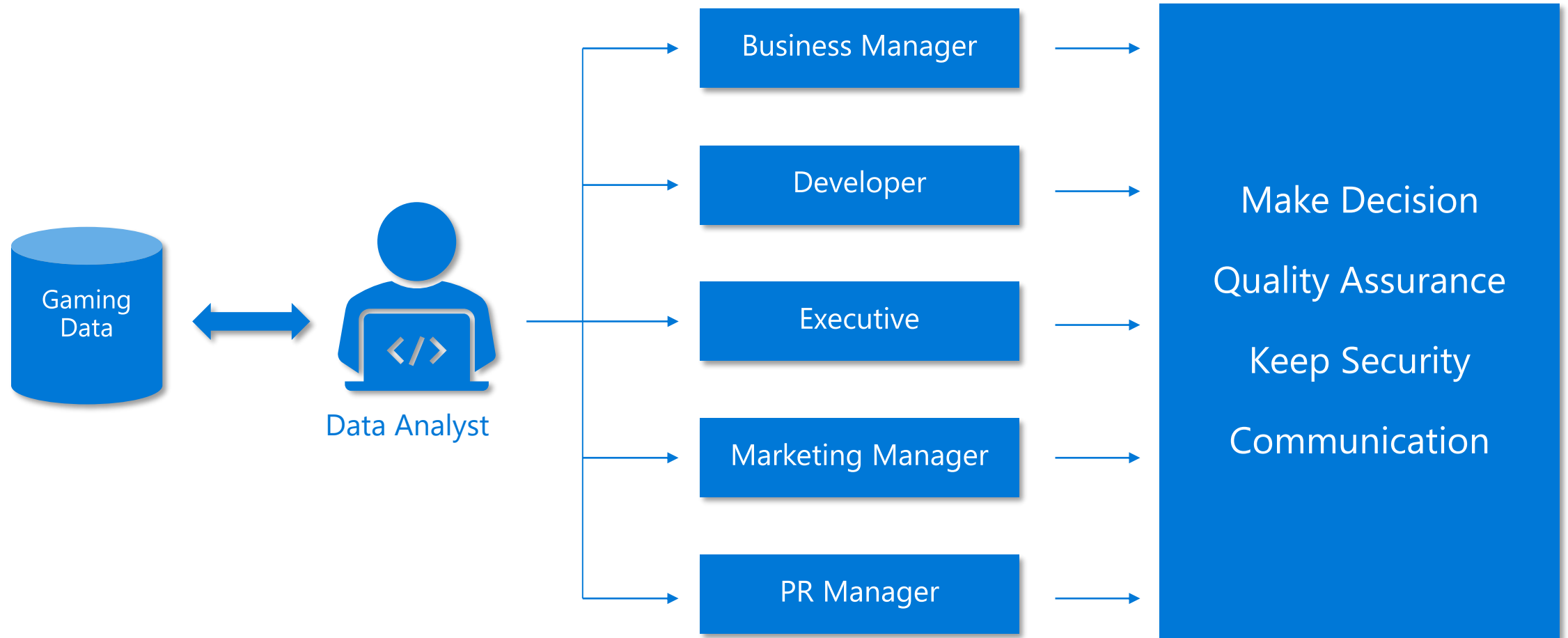


ML Modeling :
Concept & Hands -on



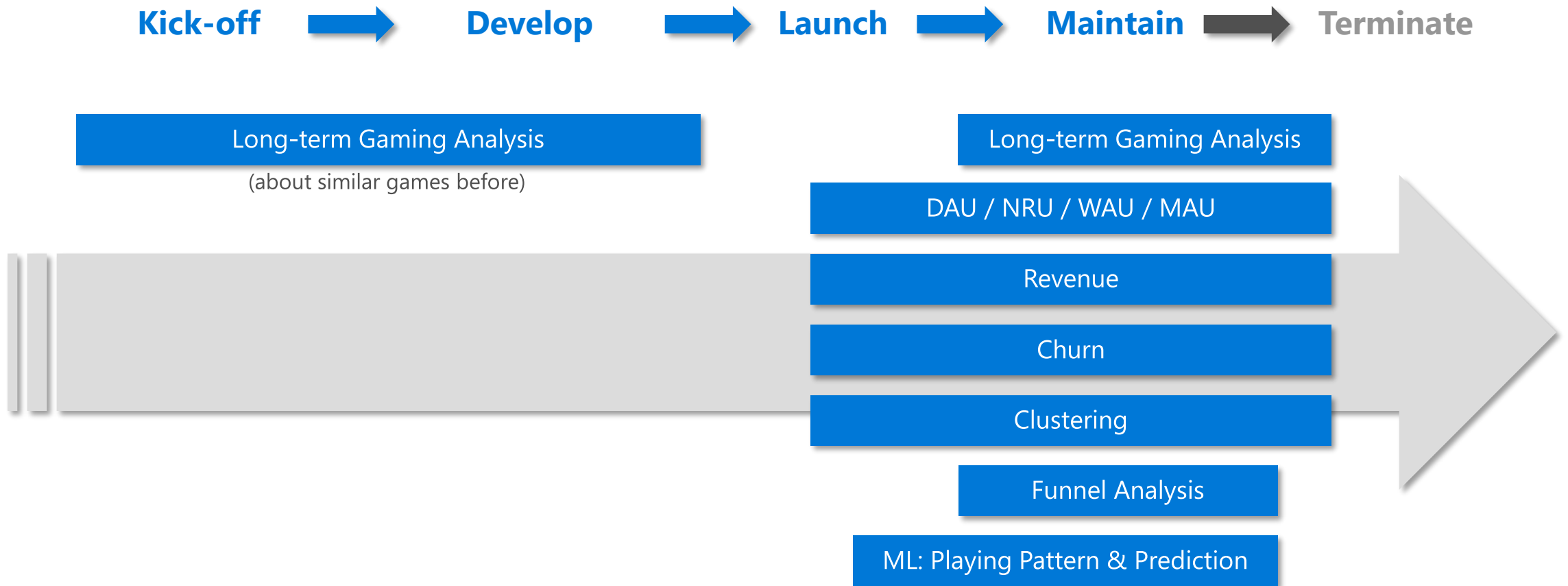
Gaming Data Analytics

Big Data Analytics in Gaming



Big Data Analytics in Gaming

Gaming: Life cycle



Gaming Data Analytics



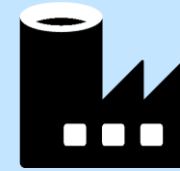
Collaboration



Integration



Visualization



ETL



Query
languages



Data
management



Easy to use



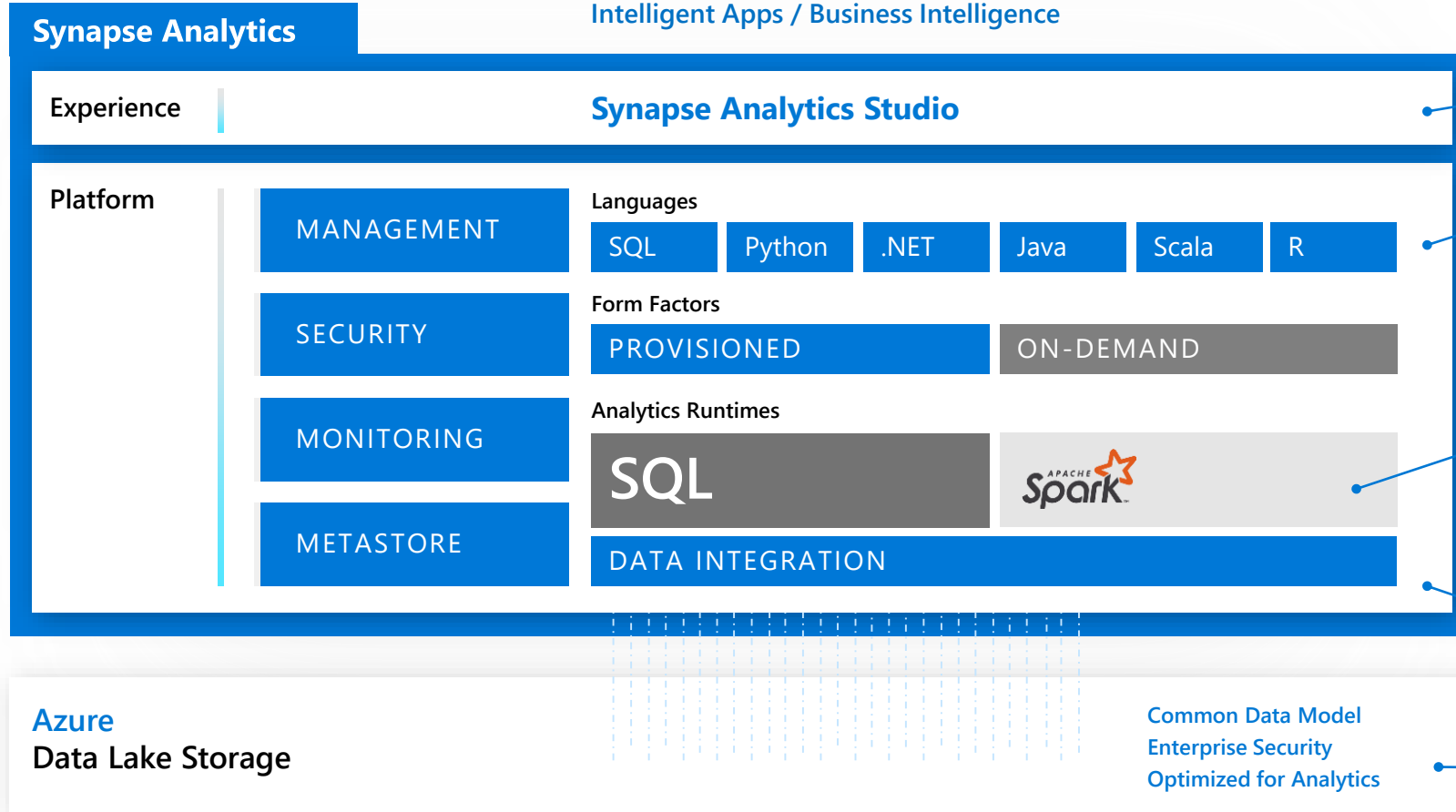
On-time

Azure Synapse Analytics

Azure Synapse Analytics

가장 빠르게 인사이트를 얻을 수 있는 무제한 분석 서비스

Artificial Intelligence / Machine Learning / Internet of Things
Intelligent Apps / Business Intelligence



Flexible한 데이터 분석 환경 제공

SaaS 개발 경험 제공
: code free / code first

다양한 분석 워크로드를 지원하는
쿼리 언어

통합 런타임 및 서버리스 환경 지원
SQL Analytics T-SQL 기반
배치/스트리밍 프로세스
Apache Spark Python, Scala, R
and .NET 기반 빅데이터 분석

관리, 보안, 모니터링,
메타스토어를 위한 통합 플랫폼
지원

Data Lake / Data model

Studio

데이터 엔지니어 / 데이터 사이언티스트 / 데이터 분석가 / IT 기술자의 데이터 분석 협업 환경


Microsoft Azure | Synapse Analytics | internalsandboxwe5


prlangad@microsoft.com
MICROSOFT


Home
Data
Develop
Orchestrate
Monitor
Manage


Synapse workspace
internalsandboxwe5

New

**Ingest**
Use the copy data tool to import data once or on a schedule.

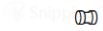
**Explore**
Learn how to navigate and interact with your data.

**Analyze**
Learn how to use SQL or Spark to get insights from your data.

**Visualize**
Build interactive reports with integrated Power BI capabilities.

Resources

Recent Pinned

NAME	LAST OPENED BY YOU
 CopyPipeline_0313	13 minutes ago

Useful links

[Synapse Analytics overview](#)
Discover the capabilities offered by Synapse and learn how to make the most of them.

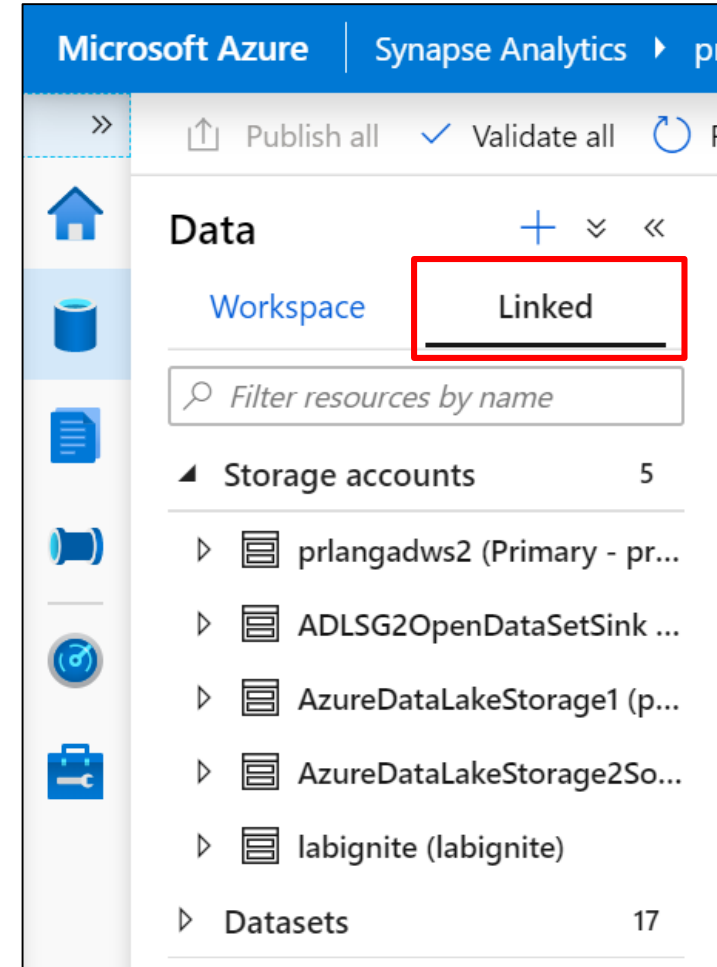
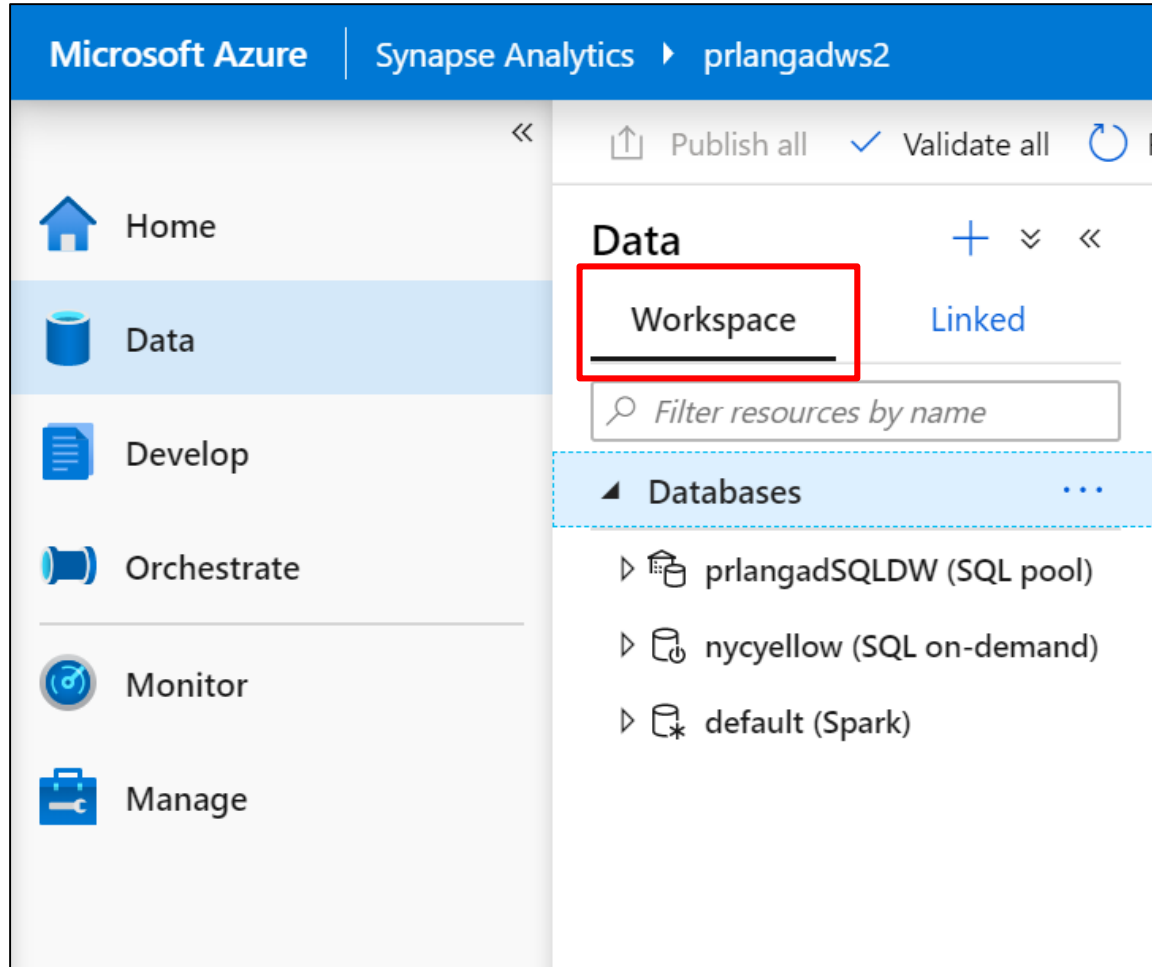
[Pricing](#)
Learn about pricing details for Synapse capabilities.

[Documentation](#)
Visit the documentation center for quickstarts, how-to guides, and references for PowerShell, APIs, etc.

[Give feedback](#)
Share your comments or suggestions with us to improve Synapse.

Data Hub

Workspace(SQL/Spark) 및 Storage 계정에 연결하여 데이터 분석



Data Hub – Storage accounts

샘플 데이터 Preview 기능 제공

The screenshot shows the Data Hub interface. On the left, under 'Storage accounts', 'internalsandboxwe (Primary)' is expanded, showing several containers. 'tempdata' is selected and highlighted in blue. On the right, the 'tempdata' storage account is open, showing a file named 'SampleCSVFile_2kb.csv'. A context menu is open over this file, with the 'Preview' option highlighted in a red box. Other options in the menu include 'New notebook', 'Copy ABFS path', 'Manage Access...', 'Rename...', 'Download', 'Delete', and 'Properties...'.

SampleCSVFile_2kb.csv

Path https://prlangaddemosa.dfs.core.windows.net/filesystem/SampleCSVFile_2kb.csv

Modified 10/29/2019, 1:30:21 PM

With column header ☒ On

USER_ID	USERNAME	FIRST_NAME	LAST_NAME	GENDER	PASSWOR
1	rogers63	david	john	Female	e6a33eee18
2	mike28	rogers	paul	Male	2e7dc6b8a1
3	rivera92	david	john	Male	1c3a8e03f4
4	ross95	maria	sanders	Male	62f0a68a41
5	paul85	morris	miller	Female	61bd060b0
6	smith34	daniel	michael	Female	7055b3d9f5
7	james84	sanders	paul	Female	b7f72d6eb9
8	daniel53	mark	mike	Male	299cbf7171
9	brooks80	morgan	maria	Female	aa736a35dc
10	morgan65	paul	miller	Female	a28dca31f5
NULL	NULL	NULL	NULL	NULL	NULL

OK

Data Hub – Storage accounts

SQL 스크립트 및 Notebook에서 간단한 쿼리 set 제공 (T-SQL 또는 PySpark 자동 생성)

← → ∨ ↑ synapsecontainer > opendataset > ISDWeathercurated

NAME
_SUCCESS
part-00000-7bc127b8-56a2-443f-b6f9-18175e9fdad7-c000.snappy.parquet
part-00000-e3d623d5-1ecd-4953-bcf9-ad973c404bd1-c000.snappy.parquet
part-00001-7bc127b8-56a2-443f-b6f9-18175e9fdad7-c000.snappy.parquet
part-00001-e3d623d5-1ecd-4953-bcf9-ad973c404bd1-c000.snappy.parquet
part-00002-7bc127b8-56a2-443f-b6f9-18175e9fdad7-c000.snappy.parquet
part-00002-e3d623d5-1ecd-4953-bcf9-ad973c404bd1-c000.snappy.parquet
part-00003-7bc127b8-56a2-443f-b6f9-18175e9fdad7-c000.snappy.parquet
part-00003-e3d623d5-1ecd-4953-bcf9-ad973c404bd1-c000.snappy.parquet
part-00004-7bc127b8-56a2-443f-b6f9-18175e9fdad7-c000.snappy.parquet
part-00004-e3d623d5-1ecd-4953-bcf9-ad973c404bd1-c000.snappy.parquet
part-00005-7bc127b8-56a2-443f-b6f9-18175e9fdad7-c000.snappy.parquet
part-00005-e3d623d5-1ecd-4953-bcf9-ad973c404bd1-c000.snappy.parquet
part-00006-7bc127b8-56a2-443f-b6f9-18175e9fdad7-c000.snappy.parquet
part-00006-e3d623d5-1ecd-4953-bcf9-ad973c404bd1-c000.snappy.parquet

New SQL script and open in new tab

New SQL script > Select TOP 100 rows

New notebook > Create external table

Copy ABFSS path

Manage Access...

Rename...

Download

Delete

Properties...

© Microsoft Corporation

opendataset SQL script 2

Run Publish Query plan Connect to SQL on-demand Use database master

```
1 SELECT
2   TOP 100 *
3 FROM
4   OPENROWSET(
5     BULK 'https://internalsandboxwe.dfs.core.windows.net/opendataset/isdweatherdatacontainer/ISDWeathercurated'
6     FORMAT='PARQUET'
7   ) AS [r];
```

USAF	WBAN	DATETIME	LATITUDE	LONGITUDE	ELEVATION
999999	04222	2019-05-01T00:2...	40.651	-122.607	432
999999	04222	2019-05-01T01:3...	40.651	-122.607	432
999999	04222	2019-05-01T02:4...	40.651	-122.607	432
999999	04222	2019-05-01T03:5...	40.651	-122.607	432
999999	04222	2019-05-01T05:0...	40.651	-122.607	432
999999	04222	2019-05-01T06:1...	40.651	-122.607	432

+ Cell ∨ Run all Undo Publish Attach to SparkPoolDef Language PySpark (Python)

Cell 1

```
[3] 1 %%pyspark
2   data_path = spark.read.load('abfss://opendataset@internalsandboxwe.dfs.core.windows.net/isdweathercurated')
3   data_path.show(100)
```

usaf	wban	datetime	latitude	longitude	elevation	windAngle	windSpeed	temperature	seaLv1Pressure	cloudCoverage	presentWeatherIndicator	pastWeatherIndicator	precipTime	precipDepth	snowDepth	stationName	countryOrRegion	year	day	month
999999	04222	2019-05-01 00:00:00	40.651	-122.607	432.0	null	2.1	19.8	null	null	null	null	null	null	null	US 2019	1	5	1	
999999	04222	2019-05-01 00:05:00	40.651	-122.607	432.0	null	null	19.7	null	null	null	null	null	null	null	US 2019	1	5	1	
999999	04222	2019-05-01 00:10:00	40.651	-122.607	432.0	null	null	19.5	null	null	null	null	null	null	null	US 2019	1	5	1	
999999	04222	2019-05-01 00:15:00	40.651	-122.607	432.0	null	null	19.2	null	null	null	null	null	null	null	US 2019	1	5	1	

Develop Hub

Overview

데이터 쿼리/분석/모델링을 위한 개발 환경

Benefits

다양한 분석 언어 (T-SQL, Python, Scala, R, .NET)

데이터 손실 없이 노트북/스크립트 간 전환

코드 IntelliSense를 탑재한 안정적인 개발 환경

시각화를 통한 인사이트 제공

Microsoft Azure | Synapse Analytics | Search resources | prlangad@microsoft.com | MICROSOFT

Home | Data | **Develop** | Orchestrate | Monitor

Develop

Filter resources by name

- SQL scripts 27
- Notebooks 11
- Data flows 2
- Spark job definitions 1
- Power BI 1

Microsoft Azure | Synapse Analytics | Search resources | prlangad@microsoft.com | MICROSOFT

opendataset | Notebook 1

Cell 1

```
1 %%pyspark
2 data_path = spark.read.load('abfs://opendataset@internalstorage.dfs.core.windows.net/holidays/part-00000-bd1ab',
3 data_path.show(100)
```

Command executed in 2mins-44s 998ms by prlangad on 03-19-2020 11:31:56-458 - 07:00

Job execution Succeeded Spark 2 executors 8 cores

ID	DESCRIPTION	STATUS	STAGES TASKS	SUBMISSION TIME	DURATION
Job 0	load at NativeMethodAccessorImpl.java0	Succeeded	1/1	3/19/2020, 11:31:35 AM	6s
Job 1	showString at NativeMethodAccessorImpl.java0	Succeeded	1/1	3/19/2020, 11:31:43 AM	1s
Job 2	showString at NativeMethodAccessorImpl.java0	Succeeded	1/1	3/19/2020, 11:31:45 AM	9s

SQL script 2

Run | Publish | Query plan

```
1 -- type your sql script here, we now have intellisense
2 CREATE
```

Keyword

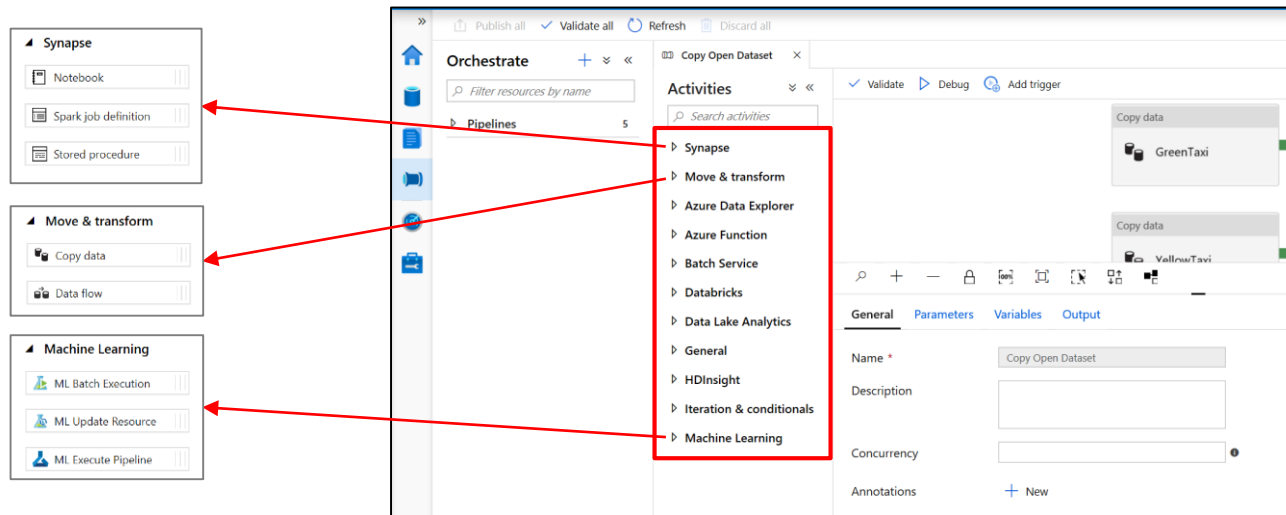
- CREATE
- CURRENT_TIMESTAMP
- CURRENT_USER

Orchestrate Hub

Pipeline 생성을 통한 ETL(데이터 적재, 변형, 로딩)

- Pipeline 상에서 수행 가능한 다양한 Activity 제공
- 90개 이상의 Connector(Linked Services)로 손쉽게 cross platform 분석 환경 구성

Pipeline 및 Activity 별 진행 상태 모니터링 (Failure 발생 시 상세 원인 내역 제공)



Pipeline runs

Time : Last week (10/24/2019 9:44 AM - 10/31/2019 9:44 AM) Time zone : Pacific Time (US & Canada) (UT... Runs : Latest runs

All status ☐ Rerun ☐ Cancel ☐ Refresh ☐ Edit columns

PIPELINE NAME	RUN START	DURATION	TRIGGERED BY	STATUS
<input type="checkbox"/> Load Data to SQLDW	10/25/2019, 3:49:42 PM	00:10:55	Manual trigger	✓ Succeeded
<input type="checkbox"/> Copy Open Dataset	10/25/2019, 2:17:54 PM	00:14:12	Manual trigger	✓ Succeeded
<input type="checkbox"/> Pipeline 1	10/24/2019, 1:23:43 PM	00:00:08	Manual trigger	✓ Succeeded

Trigger runs

Time : Last 30 days (3/17/20 11:48 PM - 4/16/20 11:48 PM) Time zone : Pacific Time (US & Canada) (UT...

All status ☐ Refresh ☐ Edit columns

Showing 1 - 1 items

TRIGGER NAME	TRIGGER TYPE	TRIGGER TIME	STATUS
TriggerCopy_csvdata100	ScheduleTrigger	4/10/20, 12:14:00 AM	✓ Succeeded

90+ Connectors out of the box

Azure (15)	Database & DW (26)		File Storage (6)	File Formats(6)	NoSQL (3)	Services and App (28)		Generic (4)
Blob storage	Amazon Redshift	Oracle	Amazon S3	AVRO	Cassandra	Amazon MWS	Oracle Service Cloud	Generic HTTP
Cosmos DB - SQL API	DB2	Phoenix	File system	Binary	Couchbase	CDS for Apps	PayPal	Generic OData
Cosmos DB - MongoDB API	Drill	PostgreSQL	FTP	Delimited Text	MongoDB	Concur	QuickBooks	Generic ODBC
Data Explorer	Google BigQuery	Presto	Google Cloud Storage	JSON		Dynamics 365	Salesforce	Generic REST
Data Lake Storage Gen1	Greenplum	SAP BW Open Hub	HDFS	ORC		Dynamics AX	SF Service Cloud	
Data Lake Storage Gen2	HBase	SAP BW via MDX	SFTP	Parquet		Dynamics CRM	SF Marketing Cloud	
Database for MariaDB	Hive	SAP HANA				Google AdWords	SAP C4C	
Database for MySQL	Apache Impala	SAP table				HubSpot	SAP ECC	
Database for PostgreSQL	Informix	Spark				Jira	ServiceNow	
File Storage	MariaDB	SQL Server				Magento	Shopify	
SQL Database	Microsoft Access	Sybase				Marketo	Square	
SQL Database MI	MySQL	Teradata				Office 365	Web table	
SQL Data Warehouse	Netezza	Vertica				Oracle Eloqua	Xero	
Search index						Oracle Responsys	Zoho	
Table storage								

Manage Hub

Overview

Management: Analytics Pool, Linked Services, Orchestration, Security

Microsoft Azureinternalsandbox

6

prlangad@microsoft.com

MICROSOFT

Home

Data

Develop

Orchestrate

Monitor

Manage

Analytics pools

SQL pools

Apache Spark pools

External connections

Linked services

Orchestration

Triggers

Integration runtimes

Security

Access control

Managed private endpoints

Publish allValidate allRefreshDiscard all

SQL pools

SQL on-demand is immediately available for your workspace. SQL pools can be configured to adapt to team or organizational requirements and constraints. [Learn more](#)

+ NewRefresh

Allow pipelines (Coming soon)

Search to filter items

Showing 1-2 of 2 items (1 On-demand, 1 User-provisioned)

NAME	TYPE	STATUS	SIZE
SQL on-demand	SQL on-demand	Online	Auto
DefSQLPool	SQL pool	Online	DW2000c

Manage – SQL pools

Overview

진행 중인 프로세스 중지 및 재시작, 스케일 변경, 태그 등록 등

Microsoft Azure | internalsandbox

>> Publish all ✓ Validate all Refresh Discard all

Analytics pools

- SQL pools
- Apache Spark pools

External connections

- Linked services

Orchestration

- Triggers
- Integration runtimes

Security

- Access control
- Managed private endpoints

SQL pools

SQL on-demand is immediately available for your workspace. SQL pools can be [Learn more](#)

+ New Refresh Allow pipelines (Coming soon)

Showing 1-2 of 2 items (1 On-demand, 1 User-provisioned)

NAME	TYPE
SQL on-demand	SQL on-demand
DefSQLPool	SQL pool

Open
Pause
Scale
AssignTags
Delete

Scale

DefSQLPool

Configure the settings that best align to the workload needs on the SQL pool. [Learn more about performance levels](#)

Performance level

DW2000c

Scale

Cancel

Manage – Apache Spark pools

Overview

진행 중인 프로세스 중지 및 재시작, 스케일 변경, 태그 등록 등

Microsoft Azure | Synapse Analytics | internalsandbox

>> Publish all Validate all Refresh Discard all

Analytics pools

- SQL pools
- Apache Spark pools**

External connections

- Linked services

Orchestration

- Triggers

Integration runtimes

- Integration runtimes

Security

- Access control
- Managed private endpoints

Apache Spark pools

Apache Spark pools can be finely tuned to run different kinds of Apache Spark applications.

+ New Refresh

Showing 1 of 1 item

NAME	SIZE
DefSparkPool	...

Open this Apache Spark pool

- Open
- Auto-pause
- Scale
- AssignTags
- Delete

Properties

Name
DefSparkPool

URL
/subscriptions/58f8824d-32b0-4825-9825-02fa6a801546/resourceGroups/SynapseSandboxRG/prov...

Creation date
03/31/2020, 4:06:19 PM

Configuration

Workspace

Packages

Upload environment config file Refresh

NAME	SIZE	DATE
requirements.txt	26B	4/8/2020, 5:24:19 PM

Close

Manage – Linked services

Overview

외부 리소스에 대한 연결 정보 확인 및 변경

Benefits

90개 이상의 pre-build connector 제공

: 데이터 저장소 혹은 Compute resource

Cross-platform data migration

Microsoft Azure | Synapse Analytics | internalsandbox

» Publish all Validate all Refresh Discard all

Analytics pools

- SQL pools
- Apache Spark pools

External connections

- Linked services

Orchestration

- Triggers
- Integration runtimes

Security

- Access control
- Managed private endpoints

Linked services

Linked services are much like connection strings, which define the connection resources.

+ New

Showing 1 - 35 of 35 items

NAME ↑↓

- AirportsSource
- analyticsstoragejrs01
- AzureDataLakeStorage_EE
- AzureSqlSynapseDemo
- AzureSynapseDWDemo
- bigdatalabstore01
- bwinadlsgen2

New linked service

PayPal (Preview)	Phoenix	PostgreSQL
Power BI	Presto (Preview)	QuickBooks (Preview)
REST	SAP BW Open Hub	SAP BW via MDX
SAP Cloud For Customer	SAP ECC	SAP HANA

Continue Cancel

Manage – Triggers

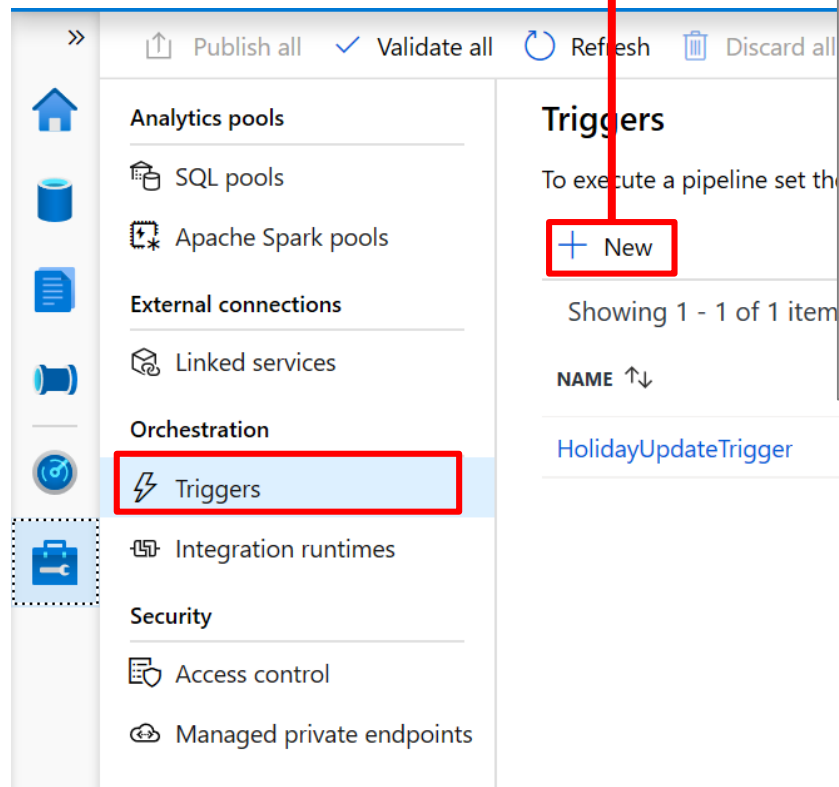
Overview

Pipeline 실행 프로세스 및 실행 주기 정의

Benefits

Create and manage

- Schedule trigger
- Tumbling window trigger
- Event trigger



New trigger

i Choose a name for your trigger. This name can be updated at any time until it is published.

Name *
Trigger 2

Description

Type *
☒ Schedule ☐ Tumbling window ☐ Event

Start Date (UTC) *
10/29/2019 9:46 PM

Recurrence *
Every 1 Minute(s)

End *
☒ No End ☐ On Date

Annotations
+ New

Activated *
☐ Yes ☒ No

OK Cancel

Manage – Access Control

Overview

RBAC(Role-Based Access Control)

: 연결된 리소스에 대한 역할별 접근 권한 설정

Benefits

- 한 개의 workspace에서 팀 공동 작업
- 생산성 향상
- SQL pool 및 Spark pool에 대한 권한 설정

Microsoft Azure | internalsandbox

» Publish all 1 Validate all Refresh Discard all

Access control

Grant access to Synapse workspace and resources by assigning a role to a user, group, service principal, or managed identity.

+ Add Refresh Remove access

Showing 1 - 3 of 3 items

NAME ↑↓	TYPE ↑↓	ROLE
Workspace admin ⓘ		
<input type="checkbox"/> Priyanka Langade Priyanka.Langade@microsoft.com	Individual	Workspace admin
<input type="checkbox"/> Synapse...		

Add role assignment

Assign a role to an AD user, group, or service principal. Each role can have up to 10 members. [Learn more](#)

Role *

Select a role

- ☐ Select all
- ☐ Workspace admin
- ☐ SQL admin
- ☐ Apache Spark admin

Apply Cancel

Demo

