

통계 기반 데이터 활용 개요

강의 내용 알아보기

강명호

데이터와 통계

13개의 질문

1. 오늘날 세계 모든 저소득 국가에서 초등학교를 나온 여성은 얼마나 될까?

- A: 20% B: 40% C: 60%

2. 세계 인구의 다수는 어디에 살까?

- A: 저소득 국가 B: 중간 소득 국가 C: 고소득 국가

3. 지난 20년간 세계 인구에서 극빈층 비율은 어떻게 바뀌었을까?

- A: 거의 2배로 늘었다. B: 거의 같다. C: 거의 절반으로 줄었다.

4. 오늘날 세계 기대 수명은 몇 세일까?

- A: 50세 B: 60세 C: 70세

5. 오늘날 전 세계 1세 아동 중 어떤 질병이든 예방접종을 받은 비율은 몇 퍼센트일까?

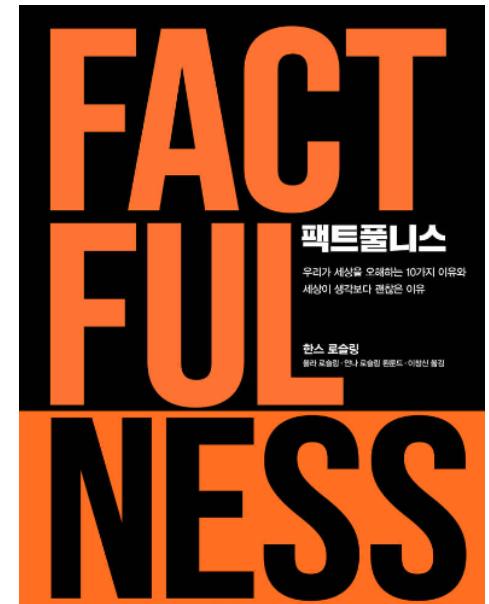
- A: 20% B: 50% C: 80%

6. 전 세계 30세 남성은 평균 10년간 학교를 다닌다. 같은 나이의 여성은 평균 몇 년간 학교를 다닐까?

- A: 9년 B: 6년 C: 3년

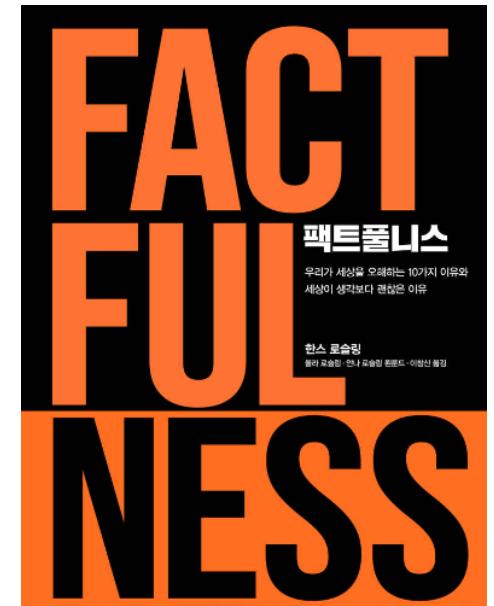
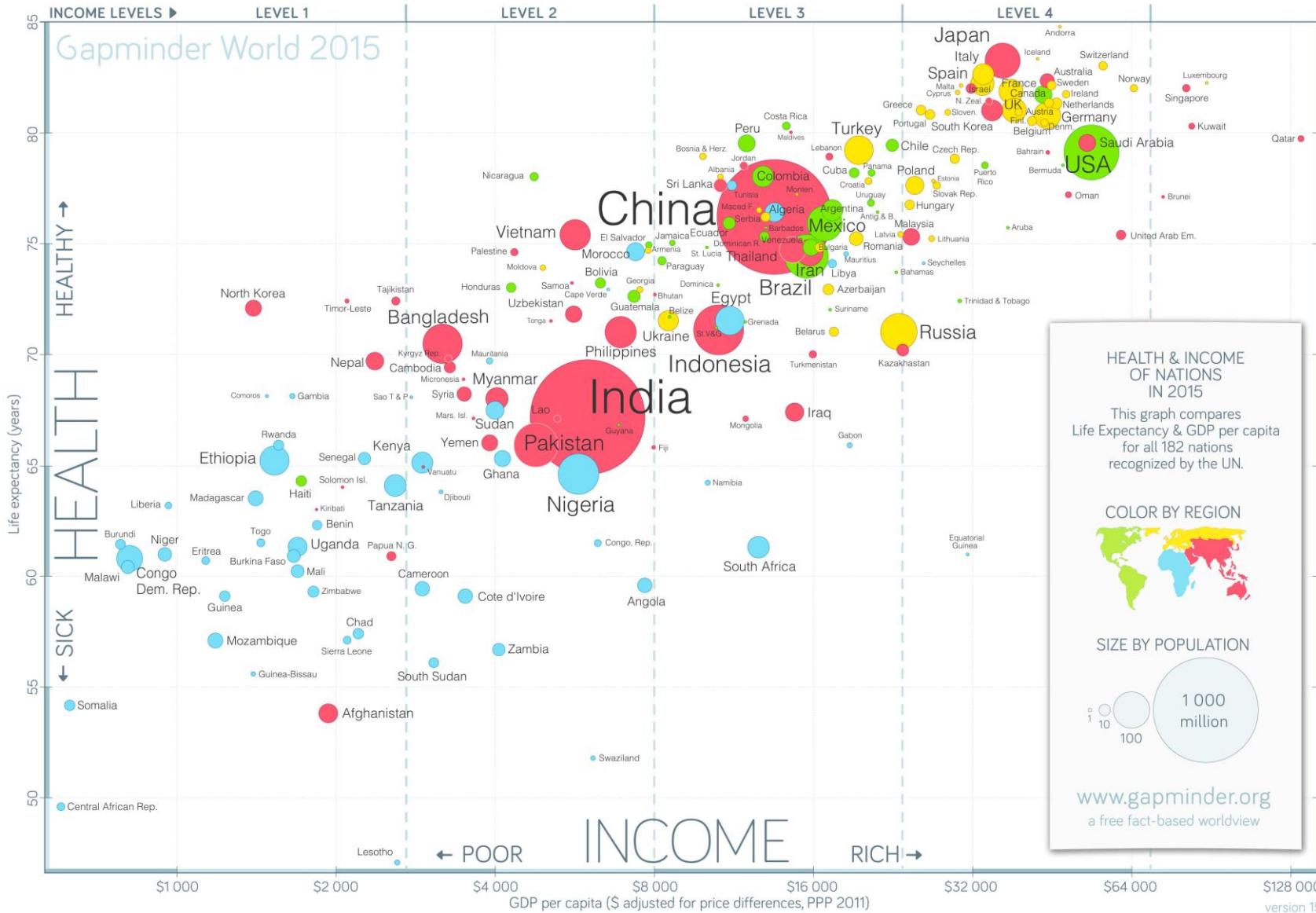
7. 세계 인구 중 어떤 식으로든 전기를 공급받는 비율은 몇 퍼센트일까?

- A: 20% B: 50% C: 80%



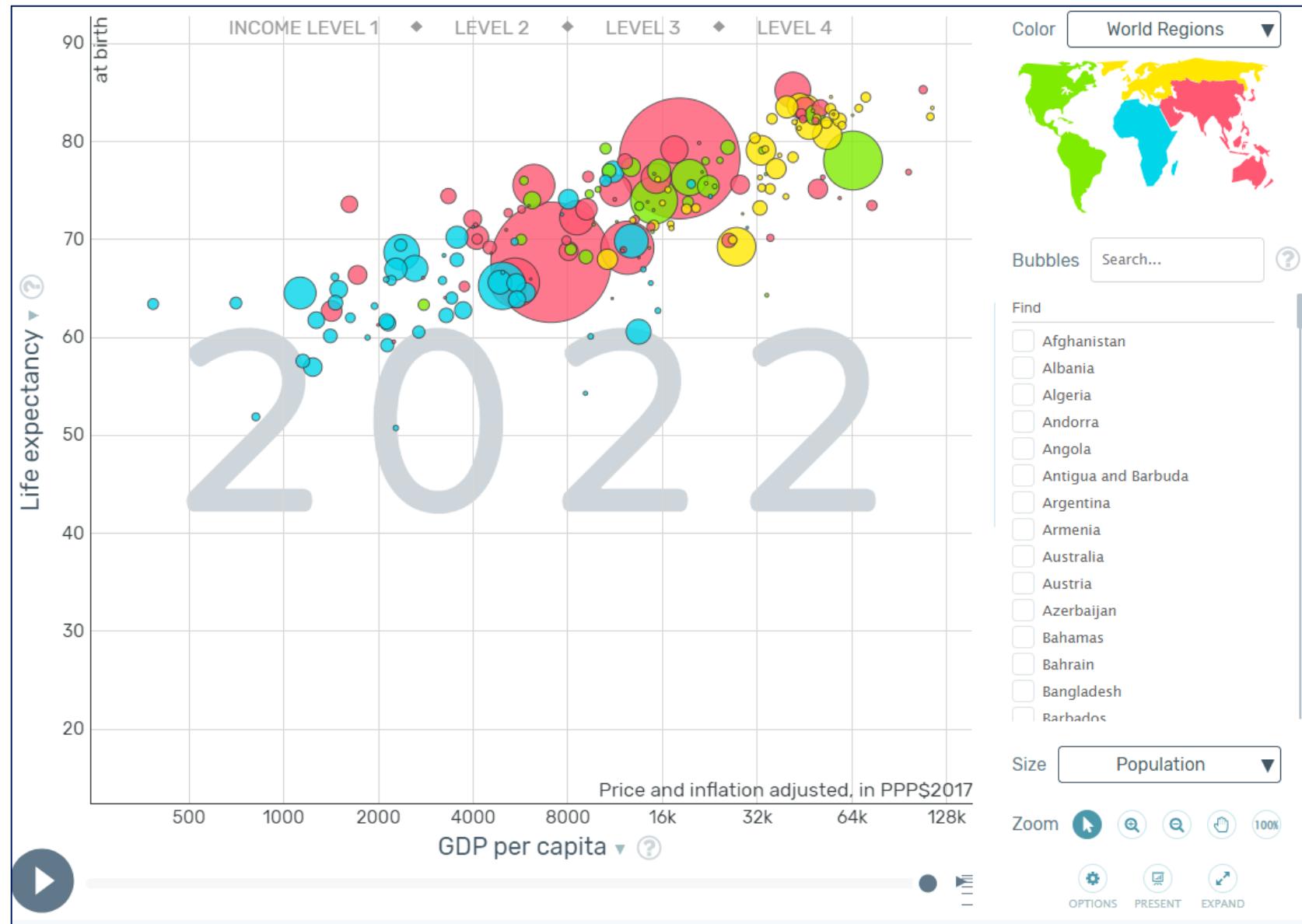
Gapminder

데이터와 통계



Gapminder

데이터와 통계



통계

통계 기본 개념

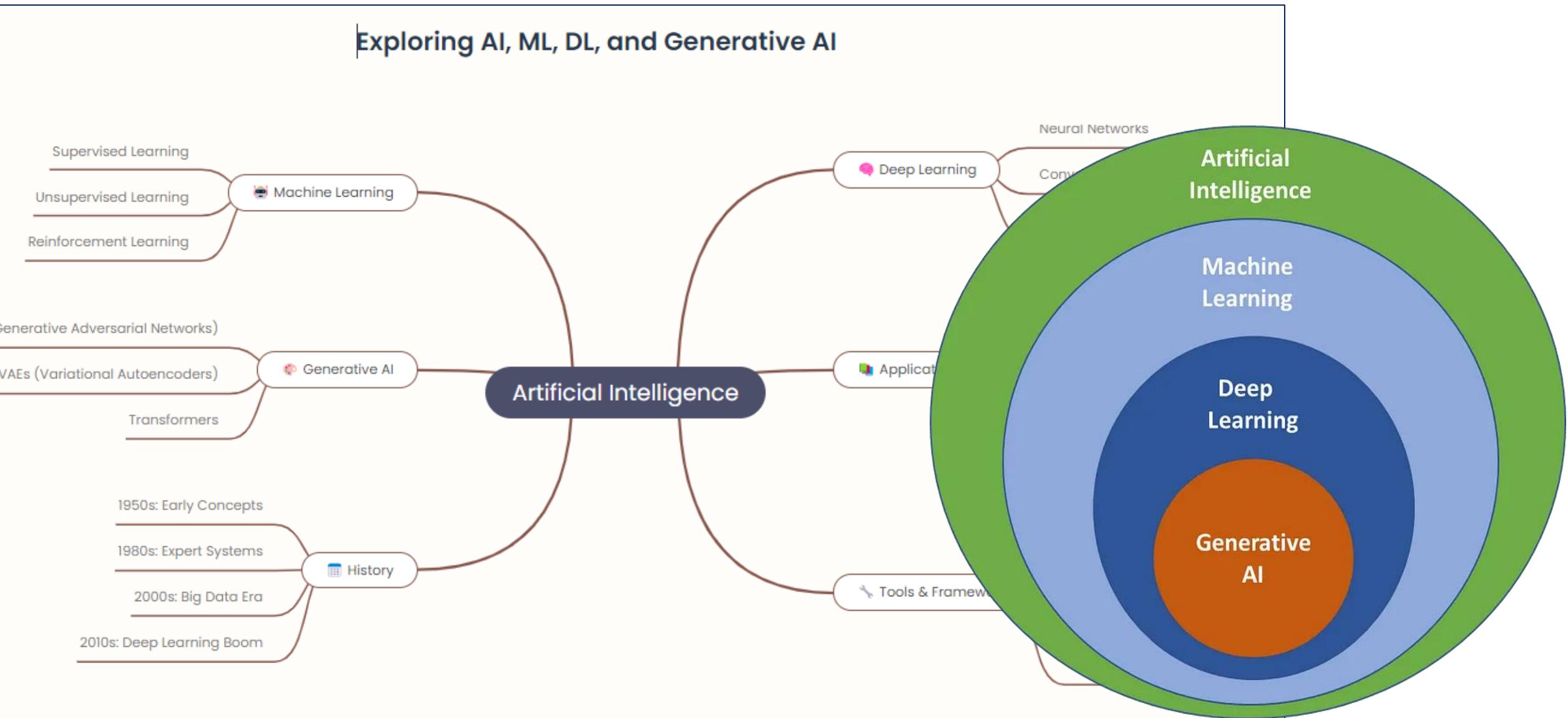
기술 통계

확률과 분포

추정과 가설 검정

상관분석

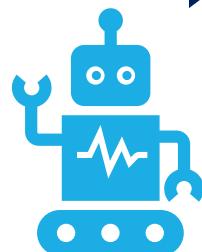
머신러닝 개요



머신러닝 개요

프로그램 기반 (기존 프로그래밍)

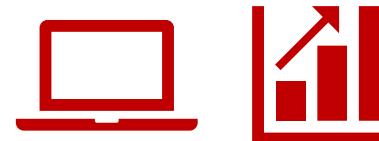
사람이 지시한
명령을 수행



- 사람이 명시적인 프로그램을 지시
- 기계는 프로그램에 기반하여 명령을 수행하고 결과 산출

데이터 기반 (머신러닝)

데이터를
통해 학습



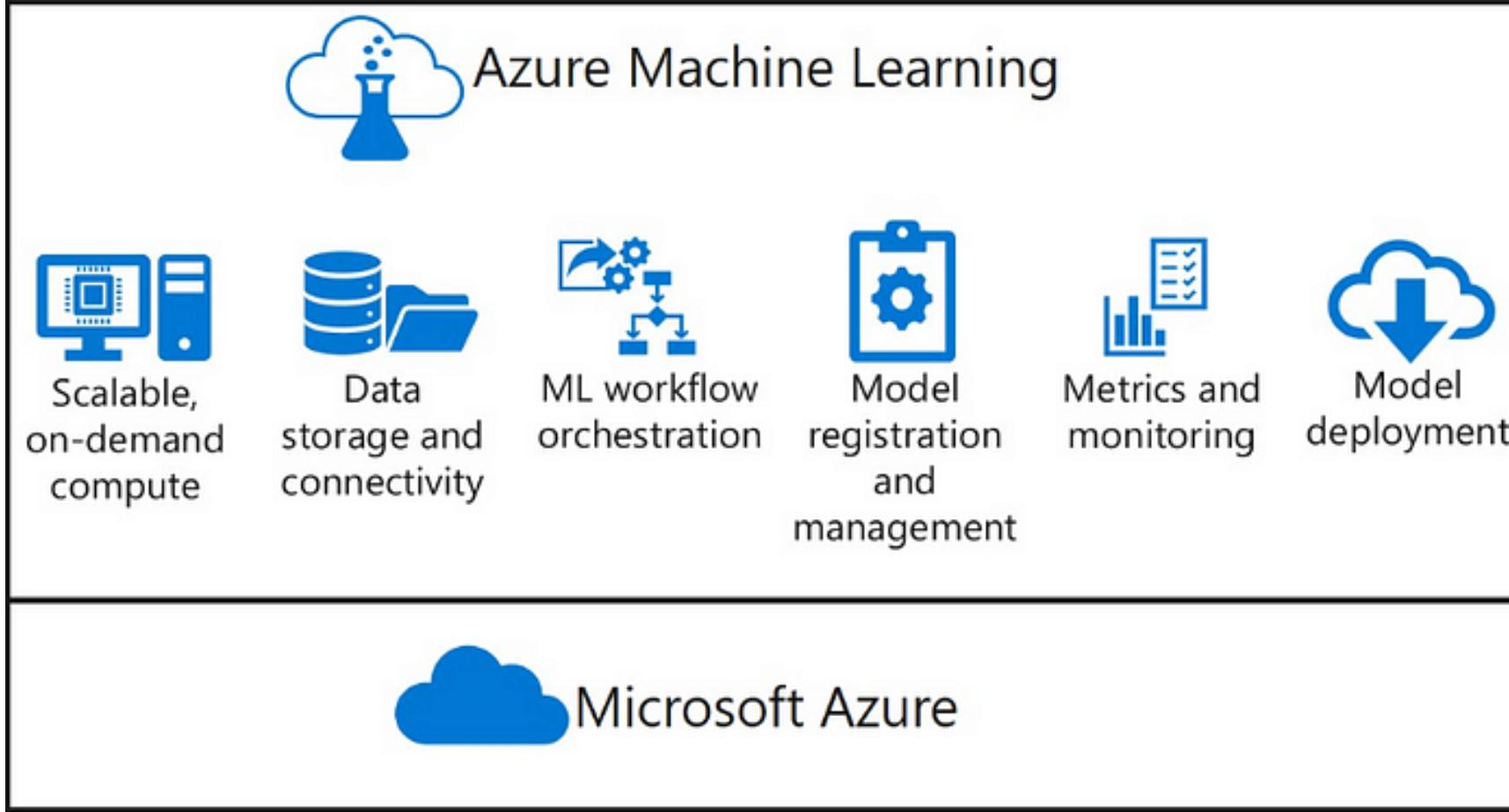
- 주어진 데이터 속에서 규칙을 학습 (규칙적인 패턴을 발견)
- 학습한 규칙을 기반으로 새로운 데이터에 대한 결과 예측

머신러닝 절차

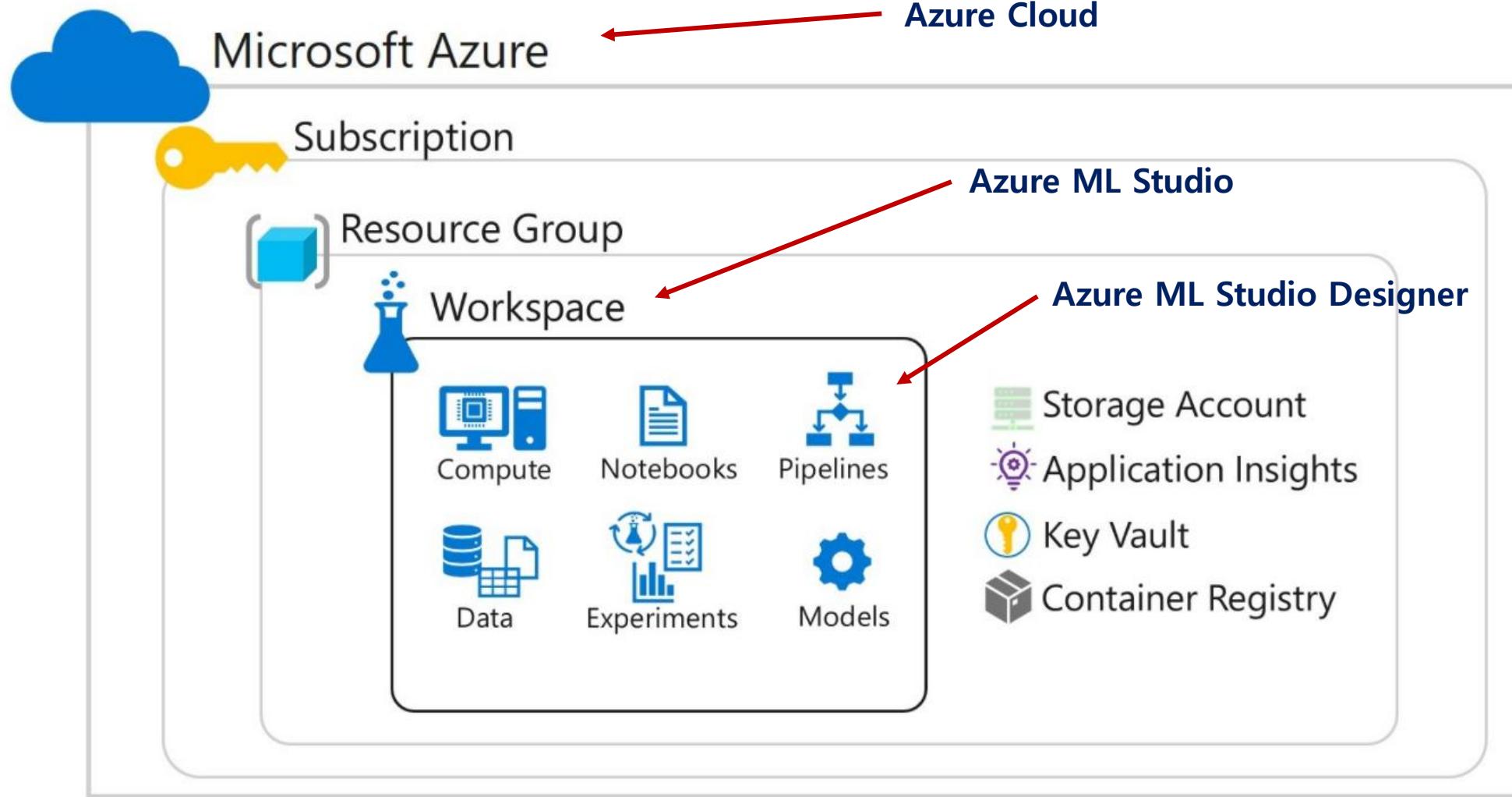


- | | | | | | | | | | | | | | | |
|--------------|------------|-----------------------------------|----------|----------|-------|------|-----------|------------------------|-------|---------|-------------|---------|-------|-------------|
| • 비즈니스 목적 정의 | • 환경/상황 분석 | • 모델링 방법 설정
(유형, 알고리즘 성능 등 검토) | • 데이터 수집 | • 데이터 이해 | • 데이터 | 레이블링 | • 데이터 전처리 | • 데이터 분리
(학습 / 테스트) | • 모델링 | 알고리즘 선택 | • 모델 학습(훈련) | • 모델 조정 | • 테스트 | • 모델링 결과 평가 |
|--------------|------------|-----------------------------------|----------|----------|-------|------|-----------|------------------------|-------|---------|-------------|---------|-------|-------------|

Azure Machine Learning



Azure Machine Learning



Azure Machine Learning – 첫 화면

Azure AI | Machine Learning Studio

고려대학교
W02-RocketLaunch

All workspaces

Home

Model catalog PREVIEW

Authoring

Notebooks

Automated ML

Designer

Prompt flow PREVIEW

Assets

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

Manage

W02-RocketLaunch

Generative AI with Prompt flow

QnA with Your Own Data Using ...
Bring Your Own Data QnA
Ask Wikipedia

Start Start Start

Generative AI models

databricks-dolly-v2-12b
Text generation

openai-whisper-large
Speech recognition

Notebook samples

View all < >

Azure Machine Learning – 데이터 등록

Authoring

- Notebooks
- Automated ML
- Designer
- Prompt flow

Assets

- Data** 
- Jobs
- Components
- Pipelines
- Environments
- Models
- Endpoints

Manage

- Compute

Data

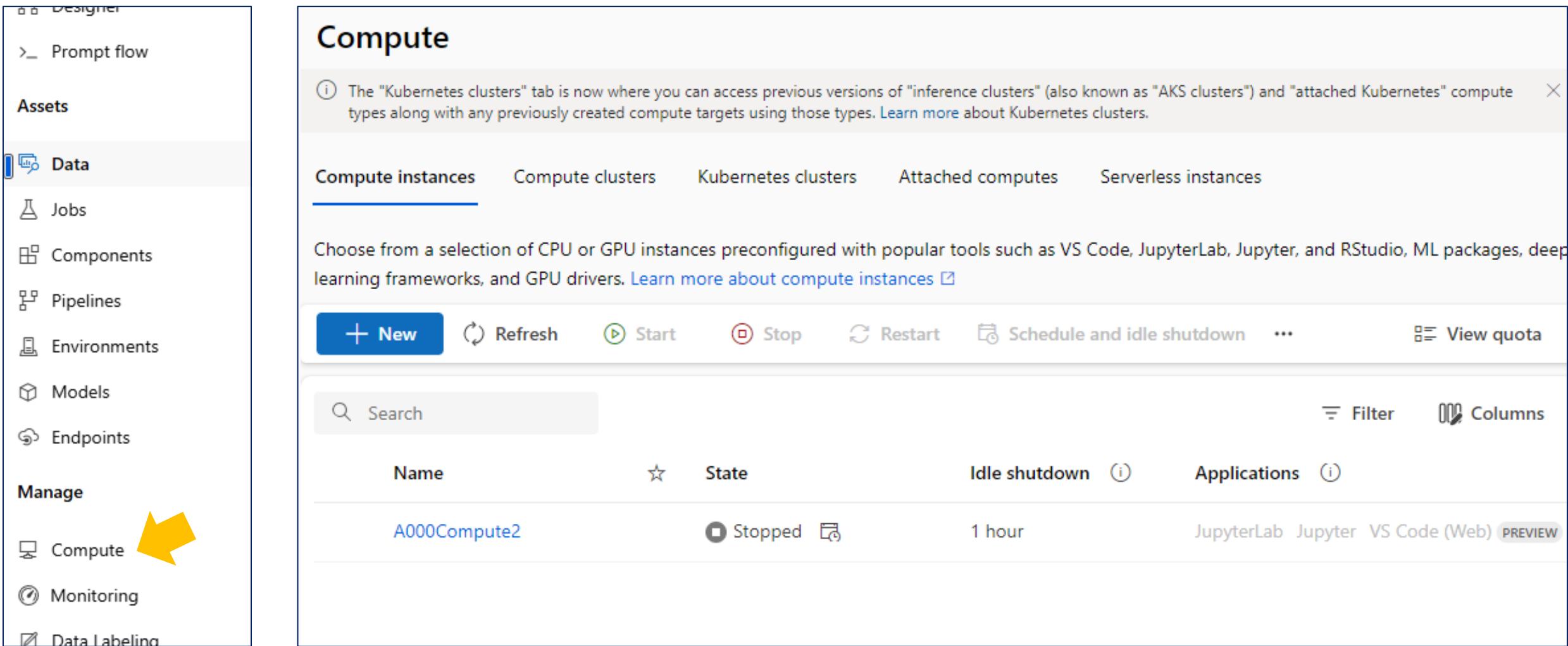
Data assets Datastores Dataset monitors PREVIEW Data import PREVIEW Data connections PREVIEW

Data assets are immutable references to your data that can be created from datastores, local files, public URLs, or Open Datasets. Data assets created with AzureML v2 cannot be deleted, but you can up-version or archive them for easy referencing and reuse in machine learning tasks. Deleting data assets created with v1 APIs will permanently delete the data asset and all metadata. [Learn more about data assets](#)

+ Create Refresh Archive Reset view Show latest version only Include archived View m

Name	Source	Version	Created on	Modified on	Type
adult_census_data	This workspace	1	Jun 19, 2024 8:41 PM	Jun 19, 2024 8:41 PM	Table

Azure Machine Learning – 컴퓨트



The screenshot shows the Azure Machine Learning Compute interface. On the left, there is a sidebar with the following items:

- Designer
- Prompt flow
- Assets
- Data
- Jobs
- Components
- Pipelines
- Environments
- Models
- Endpoints
- Manage
- Compute (highlighted with a yellow arrow)
- Monitoring
- Data Labeling

The main area is titled "Compute". It contains the following content:

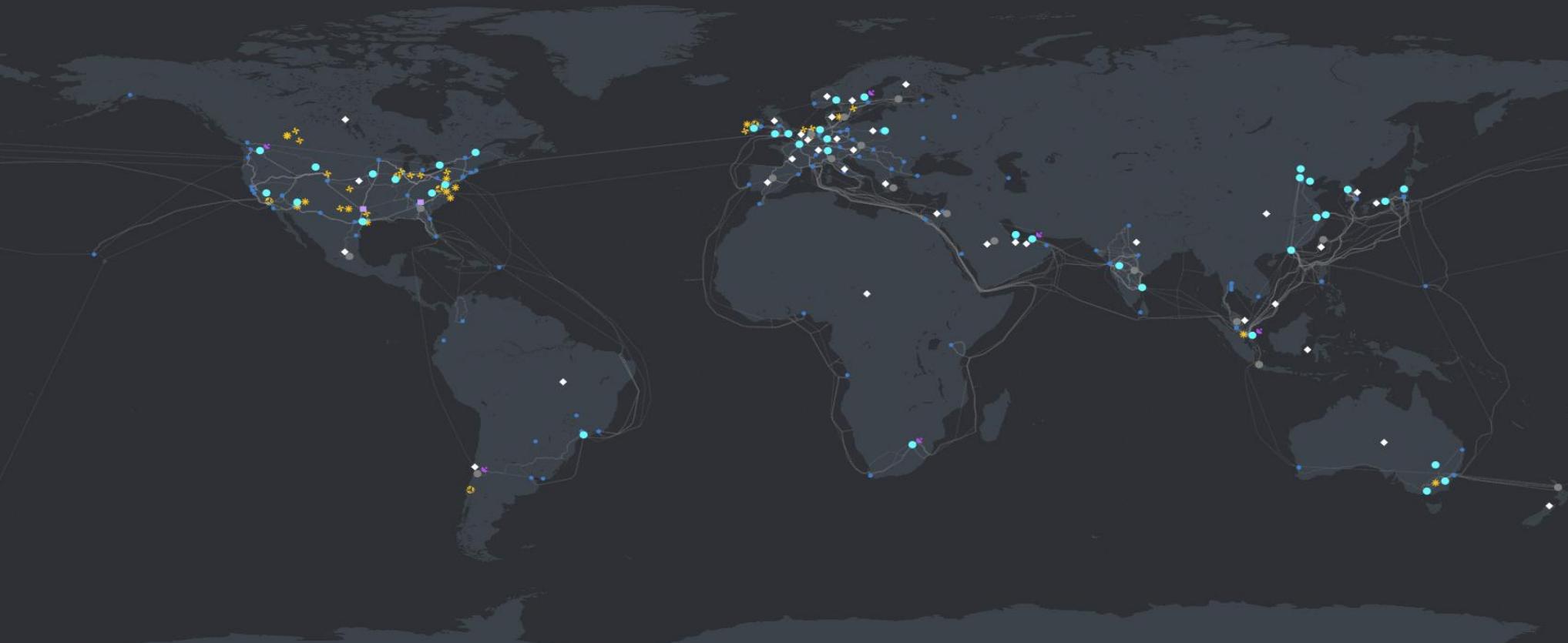
- A message: "The 'Kubernetes clusters' tab is now where you can access previous versions of 'inference clusters' (also known as 'AKS clusters') and 'attached Kubernetes' compute types along with any previously created compute targets using those types. [Learn more about Kubernetes clusters.](#)"
- Navigation tabs: Compute instances (selected), Compute clusters, Kubernetes clusters, Attached computes, Serverless instances.
- A message: "Choose from a selection of CPU or GPU instances preconfigured with popular tools such as VS Code, JupyterLab, Jupyter, and RStudio, ML packages, deep learning frameworks, and GPU drivers. [Learn more about compute instances](#)
- Actions: + New, Refresh, Start, Stop, Restart, Schedule and idle shutdown, ..., View quota.
- Search bar: Search
- Table headers: Name, State, Idle shutdown, Applications.
- Table data:

Name	State	Idle shutdown	Applications
A000Compute2	Stopped	1 hour	JupyterLab Jupyter VS Code (Web) PREVIEW
- Filter and Columns buttons.

Azure



7 News updates 



Azure Machine Learning – Designer

Authoring

- Notebooks
- Automated ML
- Designer 
- Prompt flow
- Tracing PREVIEW

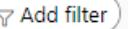
Assets

- Data
- Jobs
- Components
- Pipelines
- Environments
- Models
- Endpoints

Manage

- Compute
- Monitoring
- Data Labeling

Search by name, tags and description

Tags: All 

Data Component

95 +

Sample data (16)

Data Transformation (19)

- Add Columns Microsoft
Adds a set of columns from one dataset to another.
[\[Learn More\]](https://aka.ms/aml/add-columns) (https://aka.ms/aml/add-columns)
- azureml.Designertree  5/30/2024
- Add Rows Microsoft
Appends a set of rows from an input dataset to the end of another dataset.
[\[Learn More\]](https://aka.ms/aml/add-rows) (https://aka.ms/aml/add-rows)
- azureml.Designertree  5/30/2024
- Apply Math Operation Microsoft
Applies a mathematical operation to column values.
[\[Learn More\]](https://aka.ms/aml/apply-math-oper...) (https://aka.ms/aml/apply-math-oper...)
- azureml.Designertree  1/13/2023

Binary Classification with Feature Selection - Income Prediction 

filter_based_feature_selection
Select 5 features by ChiSquared

Two-Class Boosted Decision Tree 
vo_class_boosted_decision_tree

Untrained model  

Dataset  

Train Model 
train_model

Trained model  

Score Model 
score_model

Scored dataset  

Evaluate Model 
evaluate_model

Evaluation results 

Two-Class Boosted Decision Tree 
vo_class_boosted_decision_tree

Columns selection transformation  

Apply Transform 
apply_transform

Transform  

Select Column 
select_column

Scored dataset  

Score Model 
score_model_1

Scored dataset  

Configure & Submit

Save Pipeline interface

Create trainer mode i *

Maximum number of leaves per tree i *

Minimum number of samples per leaf no...

Learning rate i *

Number of trees constructed i *

Random number seed i

Output settings 

Input settings 

Run settings 

Azure Machine Learning – Notebooks

The screenshot shows the Azure Machine Learning studio interface. On the left, there is a sidebar with the following sections:

- Authoring**:
 - Notebooks (highlighted with a yellow arrow)
 - Automated ML
 - Designer
 - Prompt flow
 - Tracing (PREVIEW)
- Assets**:
 - Data
 - Jobs
 - Components
 - Pipelines
 - Environments
 - Models
 - Endpoints
- Manage**:
 - Compute
 - Monitoring
 - Data Labeling

The main area is titled "Notebooks" and contains a file browser and a code editor.

File Browser: Shows a tree structure with "Logs", "Users", and a folder "el01" containing a file named "Test.ipynb".

Code Editor: A notebook cell titled "Test.ipynb" is open. The code is as follows:

```
1 a = 1
2 b = 2
3 c = a + b
4 print(c)
```

The output of the cell is: [1] 3

The status bar at the top right indicates "A000Compute2 · Kernel idle CPU 0% RAM 1%" and "Last saved a few seconds ago". It also shows "Python 3.10 - SDK V2".

Azure Machine Learning – Job

Authoring

- Notebooks
- Automated ML
- Designer
- Prompt flow
- Tracing PREVIEW

Assets

- Data
- Jobs** 
- Components
- Pipelines
- Environments
- Models
- Endpoints

Manage

- Compute
- Monitoring
- Data Labeling

Jobs

All experiments All jobs All schedules

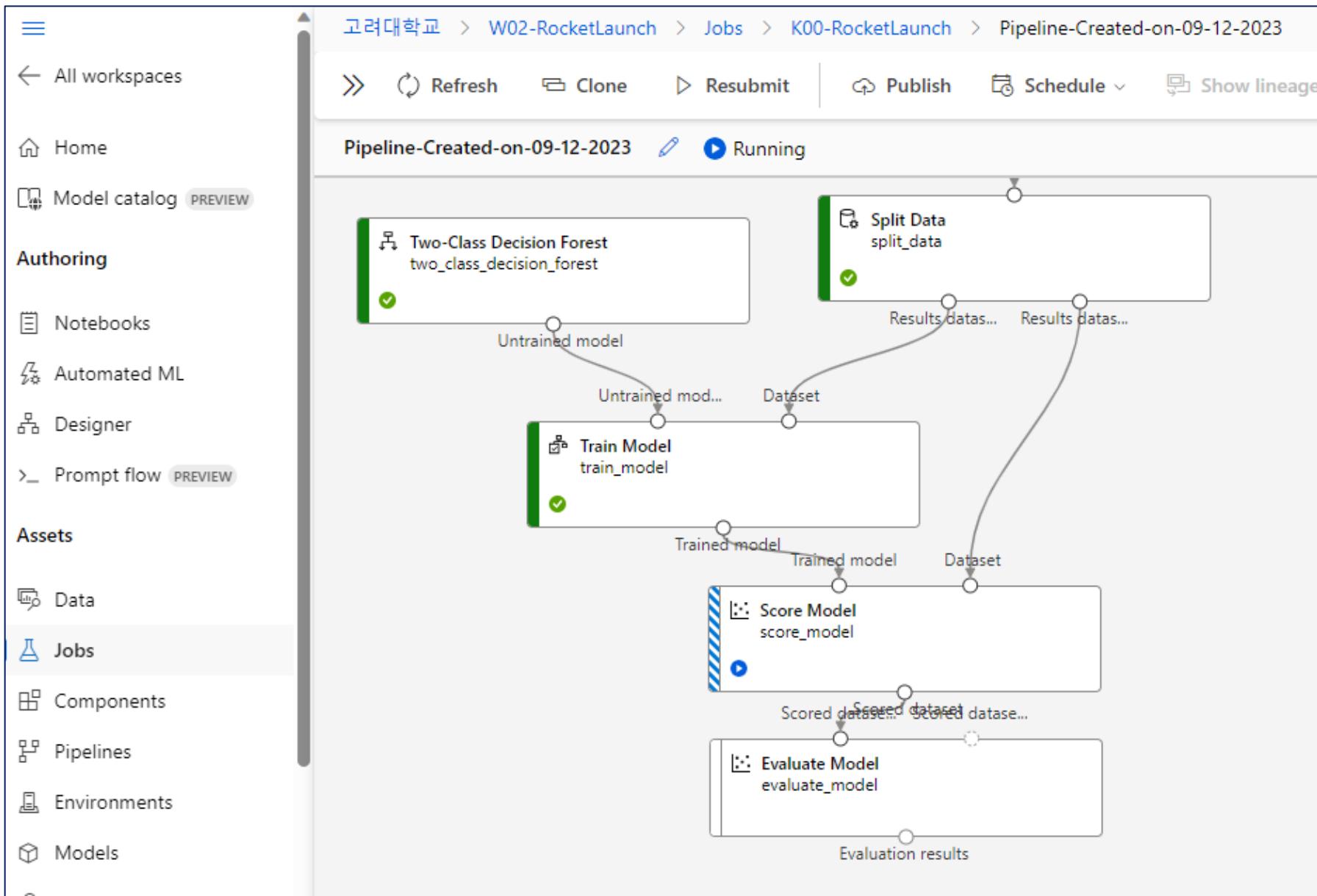
Refresh Archive experiment Reset view View archived experiments

Search 

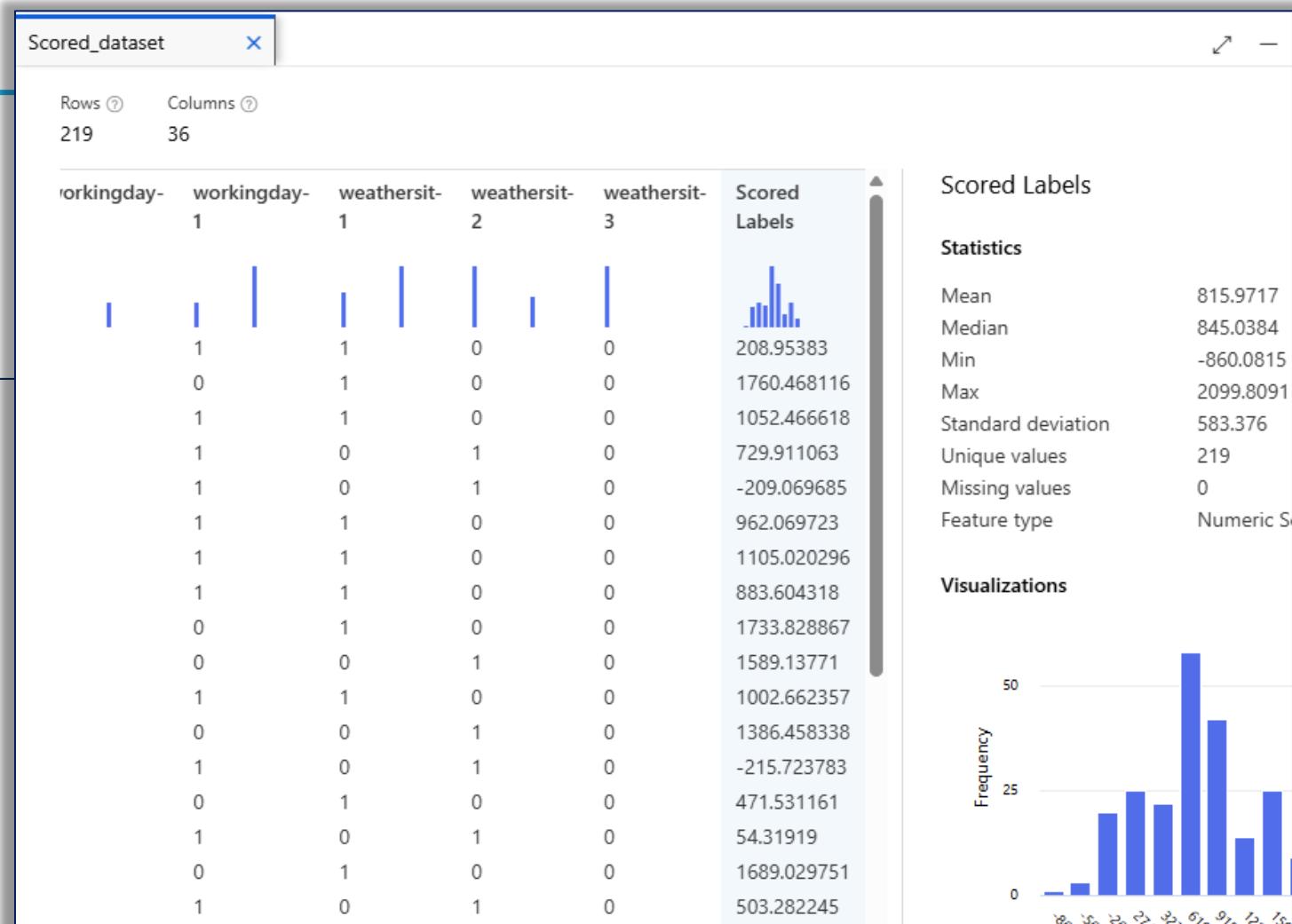
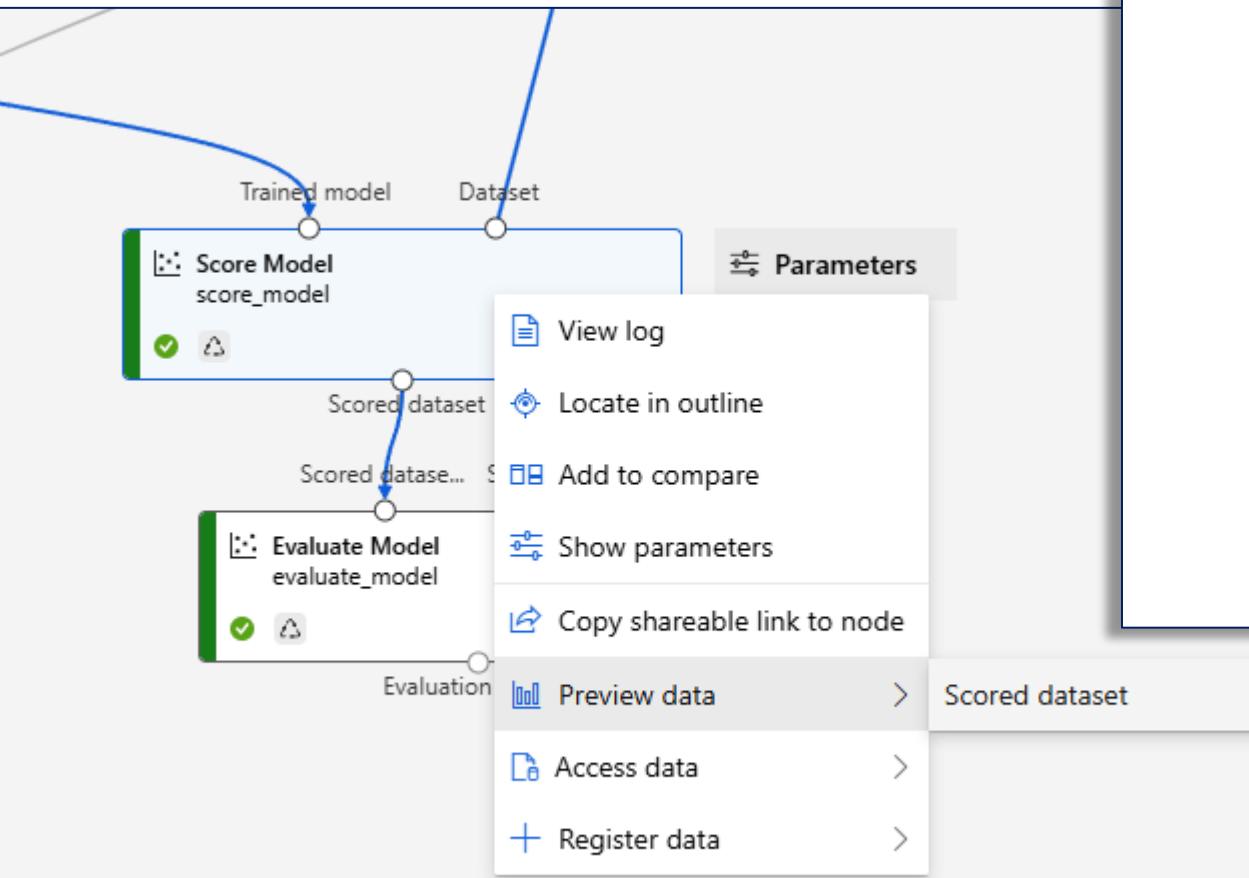
Experiment	Latest job	Last submitted
A000-CrossValidation	cross_validation	Jun 20, 2024 8:38 PM
dataset_profile	good_bird_jcltys05	Jun 20, 2024 7:34 PM
A000-incomPrediction	Pipeline-Created-on-06-19-2024	Jun 20, 2024 4:43 PM

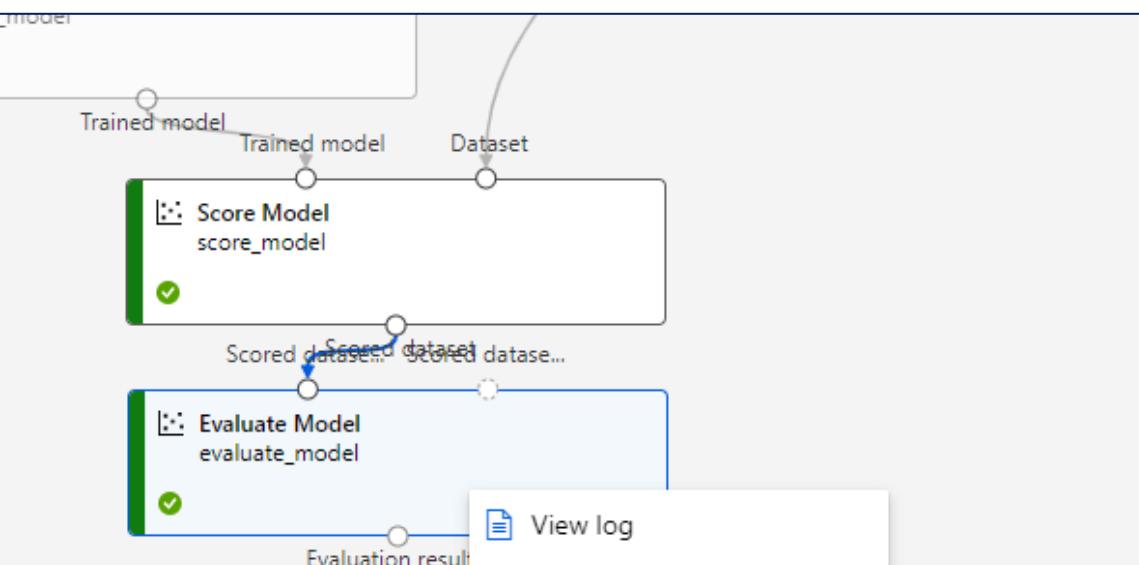
Azure Machine Learning – Job



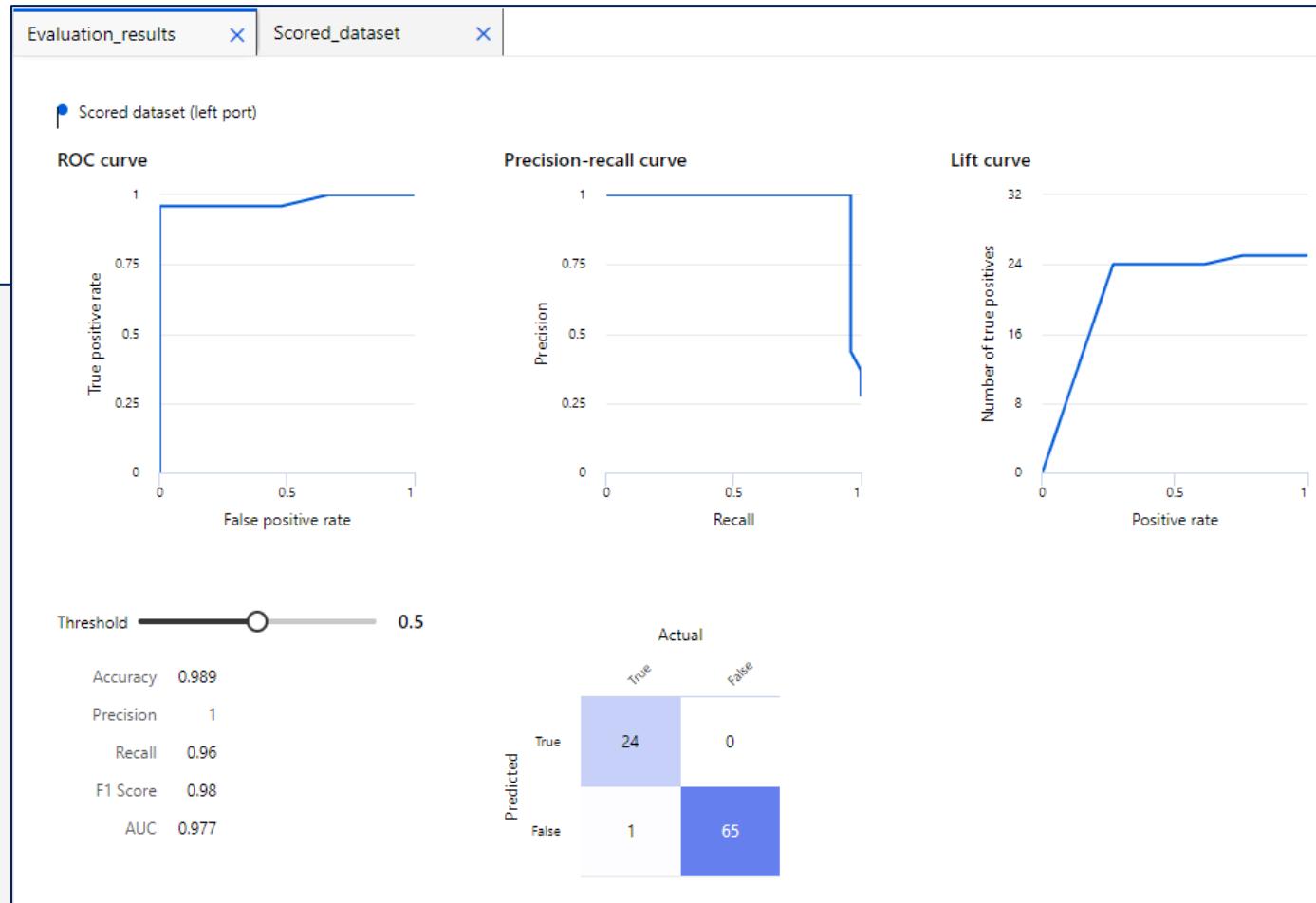
Azure Machine Learning – Job



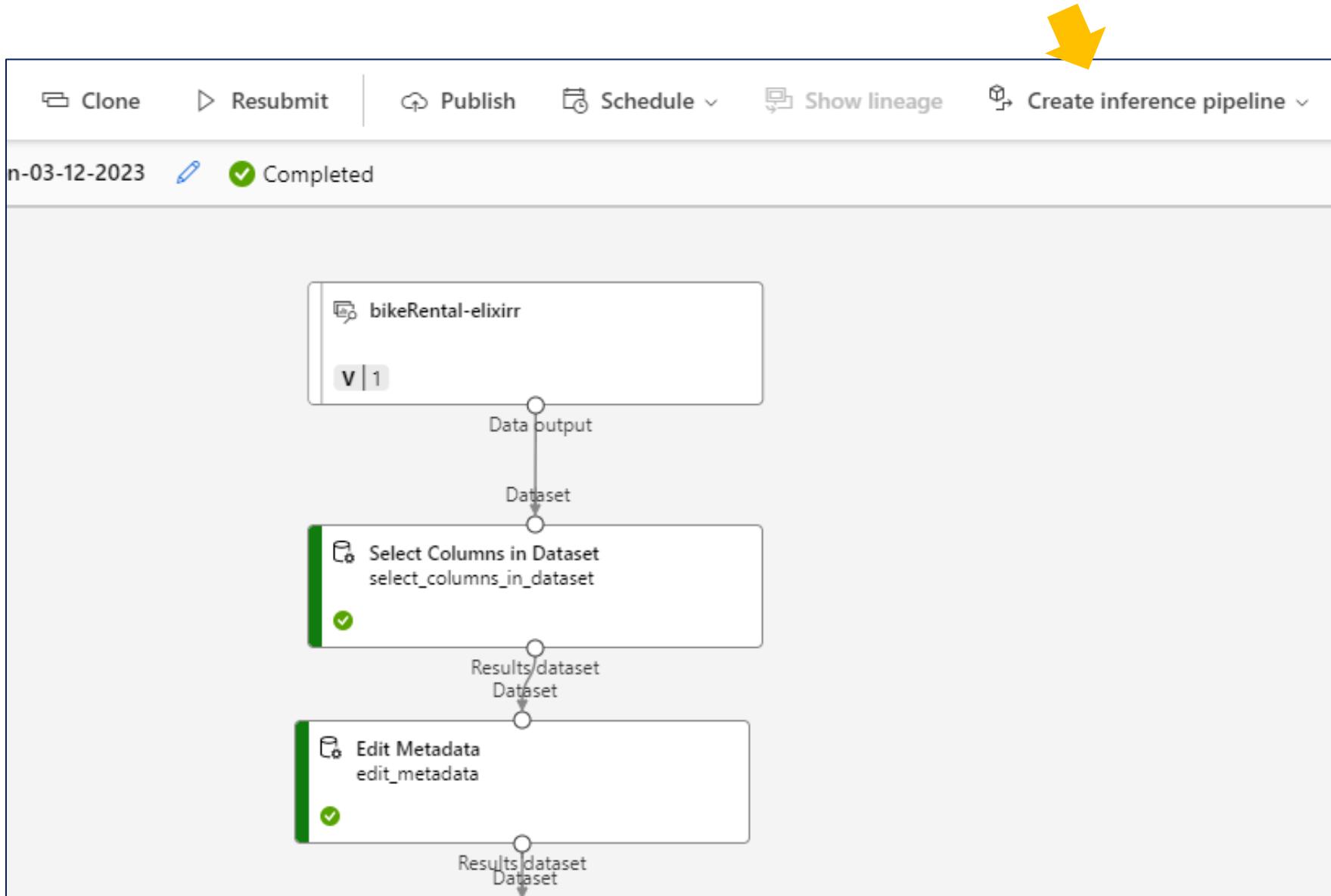
Azure Machine Learning – Job



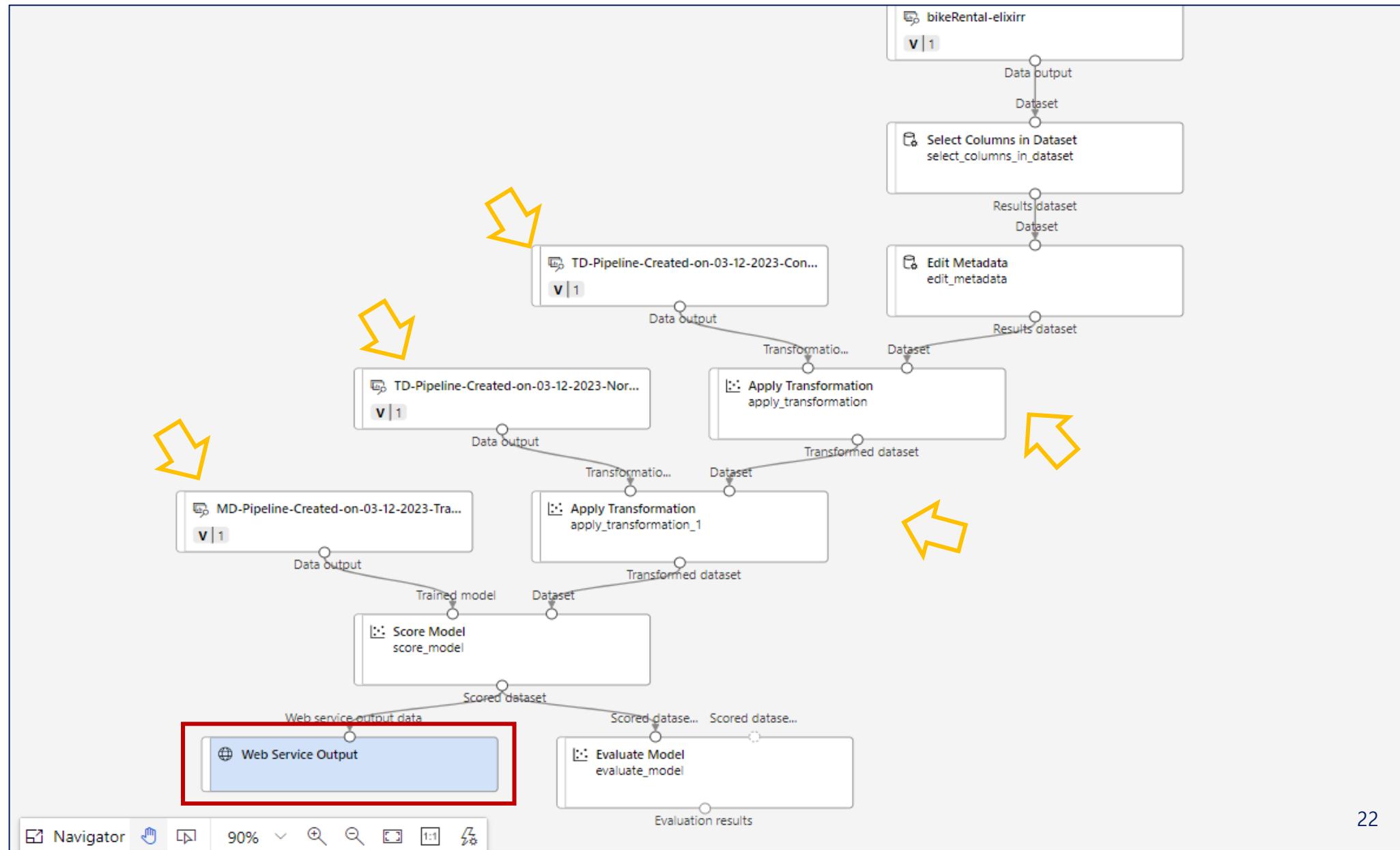
- [View log](#)
- [Locate in outline](#)
- [Copy shareable link to node](#)
- [Preview data](#)
- [Access data](#)
- [Register data](#)



Azure Machine Learning – 배포 및 유추



Azure Machine Learning – 배포 및 유추



Azure Machine Learning – 배포 및 유추

Microsoft Azure Machine Learning Studio

Microsoft > sdg-ws > Endpoints

Endpoints

New Real-time endpoints Pipeline endpoints

Refresh Disable Enable View disabled

Name ↓ Description Modified on Modified by

My_New_Pipeline	My Published Pipeline D...	11/11/2019, 3:41:03 PM	
-----------------	----------------------------	------------------------	--

< Prev Next >

Assets

Data, Jobs, Components, Pipelines, Environments, Models

Endpoints

Manage

The screenshot shows the Microsoft Azure Machine Learning Studio interface. On the left, there's a sidebar with various options like Notebooks, Automated ML, Designer, Data, Jobs, Components, Pipelines, Environments, Models, and Endpoints. The 'Endpoints' option is highlighted with a red box. In the main content area, it shows the 'Endpoints' section for the workspace 'sdg-ws'. It lists a single endpoint named 'My_New_Pipeline' with a description of 'My Published Pipeline D...', modified on 11/11/2019, 3:41:03 PM. There are tabs for 'Real-time endpoints' and 'Pipeline endpoints', with 'Pipeline endpoints' also highlighted with a red box. At the top, there are buttons for Refresh, Disable, Enable, and View disabled, along with a search bar.

Azure Machine Learning – 배포 및 유추

The screenshot shows the Azure Machine Learning Studio interface. On the left, there is a sidebar with various navigation options: Home, Model catalog (PREVIEW), Authoring, Notebooks, Automated ML, Designer, Assets, Data, Jobs, Components, Pipelines, Environments, Models, Endpoints (which is currently selected), and Manage.

The main content area displays the details for the endpoint "k00-bike-rental-inference". The title "k00-bike-rental-inference" is followed by a star icon. Below the title, there are tabs: Details, Test, Consume (which is underlined, indicating it is active), and Deployment logs.

The "Basic consumption info" section contains the REST endpoint URL: <http://2e3e01c8-d658-4a79-aa91-6d496a794899.koreacentral.azurecontainer.io/score>. There are "Regenerate" buttons next to both the Primary key and Secondary key fields.

The "Consumption option" section includes a "Consumption types" section with three buttons: Python, C#, and R. Below this, there is a code snippet:

```
1 import urllib.request
2 import json
3 import os
4 import ssl
5
6 def allowSelfSignedHttps(allowed):
7     # bypass the server certificate verification on client side
8     if allowed and not os.environ.get('PYTHONHTTPSVERIFY', '') and getattr(ssl, '_create_unverified_context', None):
```

Azure Machine Learning – Designer : 머신러닝 절차



- | | | | | |
|--------------------------------------|---------------------------------------|-------------------------------------|--|----------------------|
| • 비즈니스
목적 정의 | • 데이터 수집
• 데이터 이해
• 데이터
레이블링 | • 데이터 전처리
• 데이터 분리
(학습 / 테스트) | • 모델링
알고리즘 선택
• 모델 학습(훈련)
• 모델 조정 | • 테스트
• 모델링 결과 평가 |
| • 환경/상황 분석 | | | | |
| • 모델링 방법 설정
(유형, 알고리즘
성능 등 검토) | | | | |

Azure Machine Learning – Designer : 데이터

The screenshot shows the 'Create data asset' wizard in the Azure Machine Learning Studio. The title bar says 'Azure AI | Machine Learning Studio'. On the left, there's a sidebar with icons for Home, Data, Models, Experiments, and Assets. The main area has a breadcrumb path 'Create data asset' and two steps: '1 Data type' (highlighted with a blue circle) and '2 Data source'.

Set the name and type for your data asset

Name *
bike_rental_data

Description
자전거 렌탈 데이터 for 실습

Type * ⓘ
Tabular

Back **Next**

Azure Machine Learning – Designer : 데이터

Azure AI | Machine Learning Studio

Create data asset

1 Data type
2 Data source
3 Destination storage type
4 File or folder selection
5 Settings
6 Schema
7 Review

Select a datastore

Choose a storage type and a datastore to upload your data to in the next step. You can also create a new datastore.

Datastore type *

Azure Blob Storage

Search datastore

Name ↓	Storage name
<input checked="" type="checkbox"/> workspaceblobstore	w03bikerental4950615055
workspaceartifactstore	w03bikerental4950615055

Page 1 of 1 25/Page

Back Next

Name ↓	Storage name
<input checked="" type="checkbox"/> workspaceblobstore	w03bikerental4950615055
workspaceartifactstore	w03bikerental4950615055

Azure Machine Learning – Designer : 데이터

Create data asset

✓ Data type
✓ Data source
✓ Destination storage type
✓ File or folder selection
✓ Settings
6 Schema
7 Review

Schema
Column types are auto-detected based on the initial subset of the data and can be updated here. Values not aligning with the specified column type will fail conversion. Columns would be either null-filled or replaced with error value. Any conversions preview errors are non-blocking and you can proceed.

Search column name

Include	Column name	Type	Example values	Date format ⓘ	Properties ⓘ
<input checked="" type="checkbox"/>	Path	String		Not applicable to s...	Not applicab
<input checked="" type="checkbox"/>	day	Integer	1, 2, 3	Not applicable to s...	Not applicab
<input checked="" type="checkbox"/>	mnth	Integer	1, 1, 1	Not applicable to s...	Not applicab
<input checked="" type="checkbox"/>	year	Integer	2011, 2011, 2011	Not applicable to s...	Not applicab
<input checked="" type="checkbox"/>	season	Integer	1, 1, 1	Not applicable to s...	Not applicab
<input checked="" type="checkbox"/>	holiday	Integer	0, 0, 0	Not applicable to s...	Not applicab
<input checked="" type="checkbox"/>	weekday	Integer	6, 0, 1	Not applicable to s...	Not applicab

Back Next

Azure Machine Learning – Designer : 데이터

The screenshot shows the Azure Machine Learning Designer interface. At the top, there's a toolbar with icons for search, refresh, and navigation, followed by a menu bar with 'Undo', 'Redo', 'Validate', 'Show lineage', 'Clone', and a 'Config' button. Below the toolbar is a search bar labeled 'Search by name, tags and description' and a 'Tags : All' filter button. The main workspace displays a pipeline titled 'Pipeline-Created-on-09-19-2023'. On the left, under the 'Data' tab, there's a list of data assets: 'bike_rental_data' (Version 1, last updated 9/19/2023), 'Myeongho Kang' (고려대 자전거 렌탈 데이터), and a note about finding prebuilt sample data under the Component tab. On the right, the 'bike_rental_data' asset is selected, showing its details: ID (c881fb67-a759-49c2-83fb-f074aa467601), Data type (Tabular), Description (고려대 자전거 렌탈 데이터), Datastore name (workspaceblobstore), Relative path (UI/2023-09-19_053303_UTC/03/daily-bike-s), and Created time (Sep 19, 2023 2:46 PM). The pipeline workspace contains a single component node labeled 'bike_rental_data' with a version indicator 'V | 1' and a 'Data output' connection.

Pipeline-Created-on-09-19-2023

Save

Parameters

Data name

bike_rental_data

ID

c881fb67-a759-49c2-83fb-f074aa467601

Data type

Tabular

Description

고려대 자전거 렌탈 데이터

Datastore name

workspaceblobstore

Relative path

UI/2023-09-19_053303_UTC/03/daily-bike-s

Created time

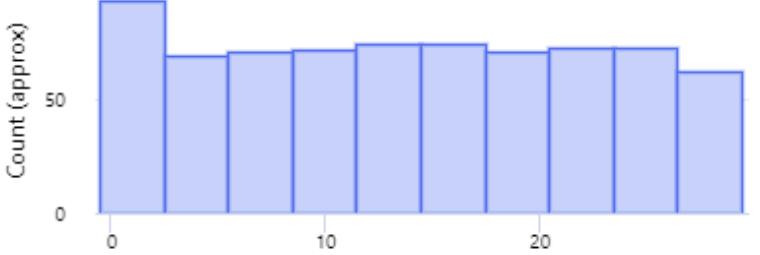
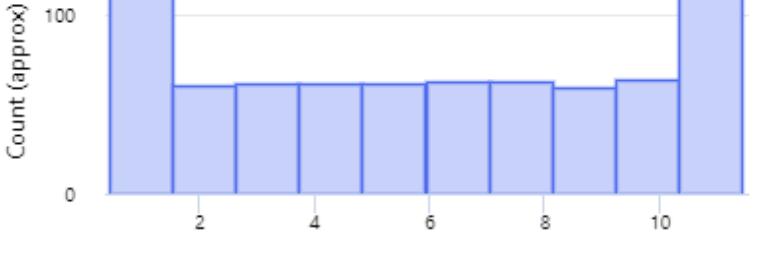
Sep 19, 2023 2:46 PM

Azure Machine Learning – Designer : 데이터

DataOutput X

Preview Profile

Number of columns: 13 Number of rows: 731

Column	Profile	Type	Min	Max	Count	Missing count	Empty count
day		Integer	1	31	731	0	0
mnth		Integer	1	12	731	0	0
year		Integer	2011	2012	731	0	0

Azure Machine Learning – Designer : 머신러닝 절차



- | | | | | |
|--------------------------------------|---------------|------------|-------------|-------------|
| • 비즈니스 목적 정의 | • 데이터 수집 | • 데이터 전처리 | • 모델링 | • 테스트 |
| • 환경/상황 분석 | • 데이터 이해 | • 데이터 분리 | 알고리즘 선택 | • 모델링 결과 평가 |
| • 모델링 방법 설정
(유형, 알고리즘
성능 등 검토) | • 데이터
레이블링 | (학습 / 테스트) | • 모델 학습(훈련) | • 모델 조정 |

Azure Machine Learning – Designer : 데이터 준비

고려대학교 > W03-bikeRental > Designer > Authoring

select columns in X

Tags : All Add filter

Data Component

3 + Most relevant

Select Columns in Dataset Microsoft Select columns to include or exclude from a dataset in an operation. [Learn More](<https://aka.ms/aml/select-columns-in-dataset>) azurerm.Designer:true 1/13/2023

Join Data Microsoft Joins two datasets on selected key columns. [Learn More](<https://aka.ms/aml/join-data>) azurerm.Designer:true 1/13/2023

Select Columns Transform Microsoft

Pipeline-Created-on-09-19-2023

Undo Redo Validate Show lineage Clone ... Configure & Submit Save Pipeline interface

bike_rental_data bike_rental_data v | 1 Data output

Select Columns in Dataset select_columns_in_dataset Results dataset

Parameters

Select Columns in Dataset

Select columns * Edit column

A value is required.

Output settings

Input settings

Run settings

Node information

Component information

```
graph TD; A[bike_rental_data] --> B[Select Columns in Dataset]; B --> C[Select Columns in Dataset]; C --> D[Results dataset]
```

Azure Machine Learning – Designer : 데이터 준비

Screenshot of the Azure Machine Learning Designer interface showing a pipeline for data preparation.

The pipeline consists of the following components:

- edit_metadata**: An **Edit Metadata** component with an output named **Results dataset**.
- convert_to_indicator_values**: A **Convert to Indicator Values** component with an output named **Indicator val...**.
- normalize_data**: A **Normalize Data** component with an output named **Transformed d...**.

A yellow arrow points to the **normalize_data** component.

Normalize Data component settings:

- Transformation method**: ZScore
- Use 0 for constant columns when checked..**: True
- Columns to transform**: temp, atemp, hum, windspeed

Other pipeline details:

- Tags**: All
- Data**: 1 item
- Component**: Most relevant
- Created**: Pipeline-Created-on-09-19-2023
- Last modified**: 1/13/2023
- Owner**: Microsoft
- Description**: Rescales numeric **data** to constrain **dataset** values to a standard range. [Learn More](https://aka.ms/aml/...)
- Tags**: azureml.Designer=true

Azure Machine Learning – Designer : 데이터 준비

The screenshot shows the Azure Machine Learning Designer interface. On the left, there is a sidebar with a search bar for "edit metadata", a "Tags : All" button, and a "Component" tab selected. Below this, there is a list of components: "Edit Metadata" (selected), "Microsoft", and "Select Columns in Dataset". The "Edit Metadata" component has a tooltip: "Edits metadata associated with columns in a dataset." and a link "[Learn More](https://aka.ms/aml/edit-metadata)". The date "1/13/2023" is also shown. A yellow arrow points to the "Edit Metadata" component in the main pipeline.

The main area displays a pipeline titled "Pipeline-Created-on-09-19-2023". The pipeline consists of three components connected by arrows:

- The first component is "bike_rental_data" (type: Dataset).
- An arrow labeled "Data output" points from "bike_rental_data" to the second component, "Select Columns in Dataset".
- The second component is "Select Columns in Dataset" (type: Dataset).
- An arrow labeled "Results dataset" points from "Select Columns in Dataset" to the third component.
- The third component is "Edit Metadata" (type: Dataset).
- An arrow labeled "Results dataset" points from the "Edit Metadata" component back to the first "Dataset" component.

On the right side, there is a "Save" button and a "Pipeline interface" button. A large "Edit Metadata" panel is open on the right, showing configuration options for the selected component. The "Column" field is required and currently empty. The "Data type" is set to "Unchanged". The "Categorical" field is also set to "Unchanged". The "Fields" field is set to "Unchanged". The "New column names" field is empty.

Azure Machine Learning – Designer : 데이터 준비

The screenshot shows the Azure Machine Learning Designer interface with the following components:

- Search Bar:** Displays "split data".
- Filter:** Shows "Tags : All" and "Add filter".
- Component View:** Shows two components:
 - Split Data:** Microsoft, Partitions the rows of a dataset into two distinct sets. [Learn More](https://aka.ms/aml/split-data). Last modified 1/13/2023.
 - Apply SQL Transformation:** Microsoft, Runs a SQLite query on input datasets to transform the data. [Learn More](https://aka.ms/aml/apply-sq...). Last modified 1/13/2023.
- Pipeline Editor:** Pipeline-Created-on-09-19-2023. The pipeline consists of the following steps:
 - Results dataset → Convert to Indicator Values (convert_to_indicator_values)
 - Convert to Indicator Values → Results dataset, Indicator values Dataset
 - Indicator values Dataset → Normalize Data (normalize_data)
 - Normalize Data → Transformed dataset, Transformation...
 - Transformed dataset → Split Data (split_data)
 - Split Data → Results dataset, Results dataset
- Split Data Component Detail:** Shows configuration for Split Data:
 - Splitting mode: Split Rows
 - Fraction of rows in the first output: 0.5
 - Randomized split: True
 - Random seed: 0
 - Stratified split: False

Azure Machine Learning – Designer : 머신러닝 절차



- | | | | | |
|--------------------------------------|---------------|------------|-------------|-------------|
| • 비즈니스 목적 정의 | • 데이터 수집 | • 데이터 전처리 | • 모델링 | • 테스트 |
| • 환경/상황 분석 | • 데이터 이해 | • 데이터 분리 | 알고리즘 선택 | • 모델링 결과 평가 |
| • 모델링 방법 설정
(유형, 알고리즘
성능 등 검토) | • 데이터
레이블링 | (학습 / 테스트) | • 모델 학습(훈련) | • 모델 조정 |

Azure Machine Learning – Designer : 모델링

Virtual machine type [\(i\)](#)

CPU GPU

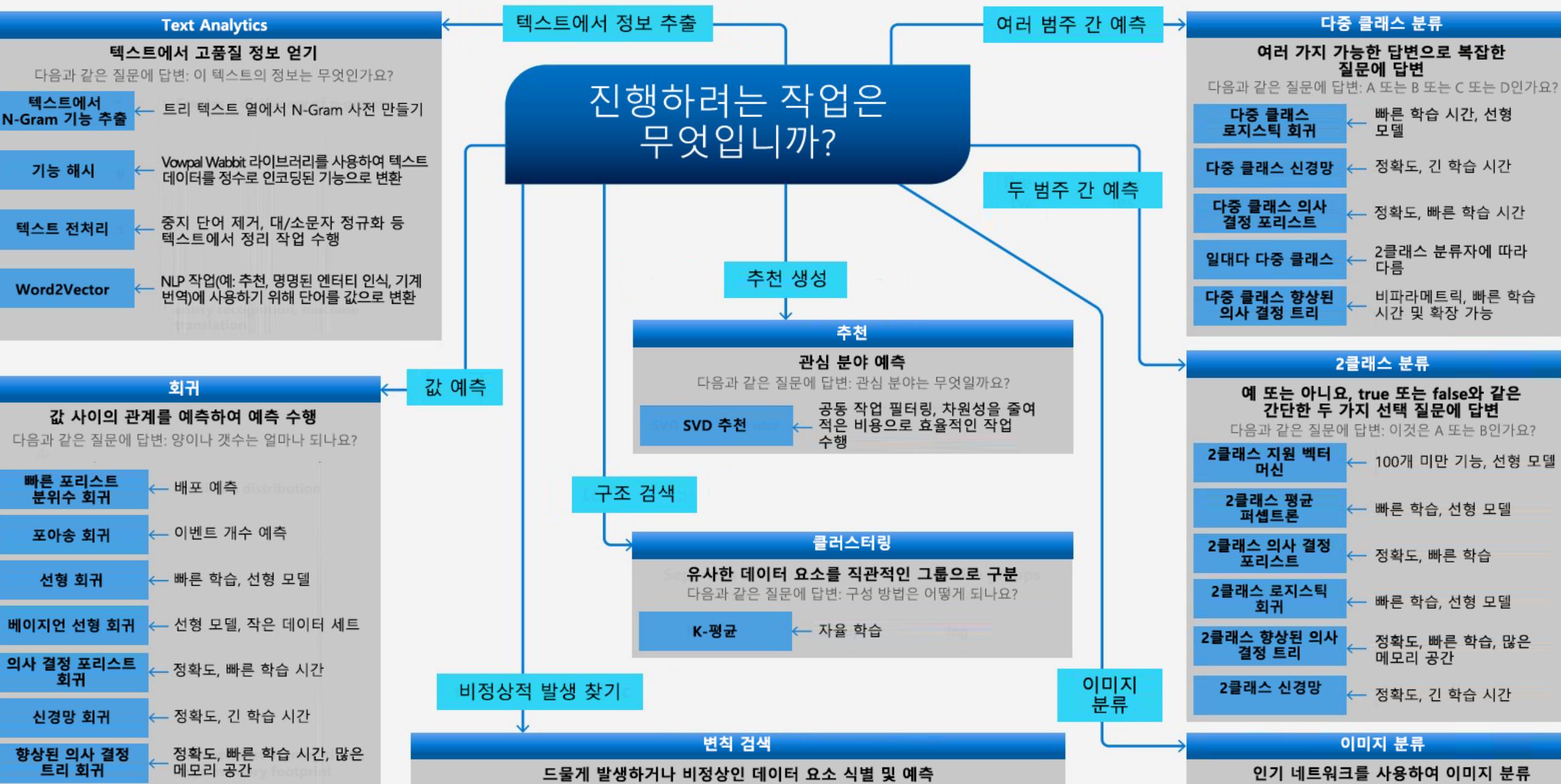
Virtual machine size [\(i\)](#)

Select from recommended options Select from all options

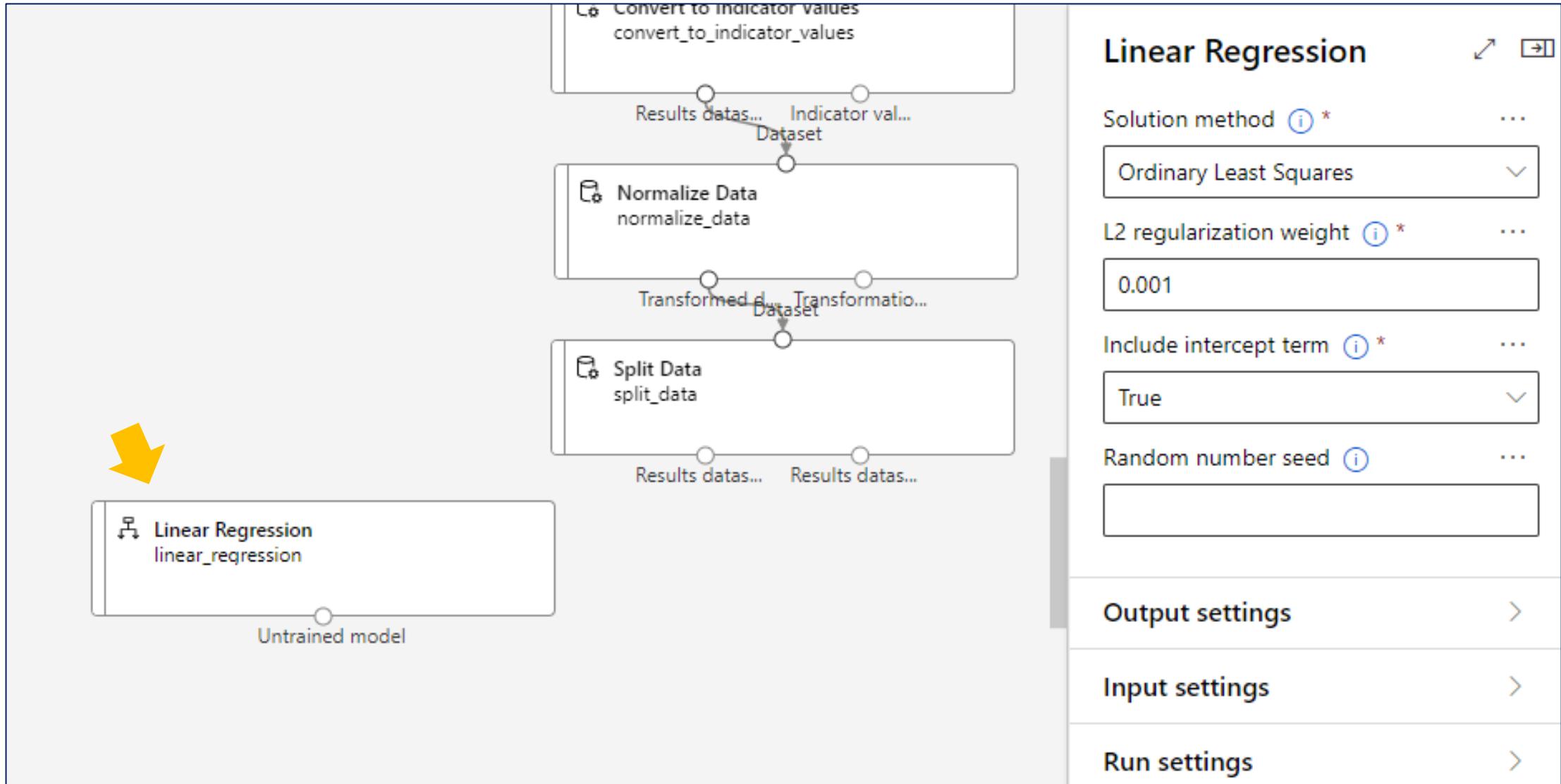
[+▽ Add filter](#) Search by VM name...

Showing 200 VM sizes | Current selection: Standard_E4ds_v4

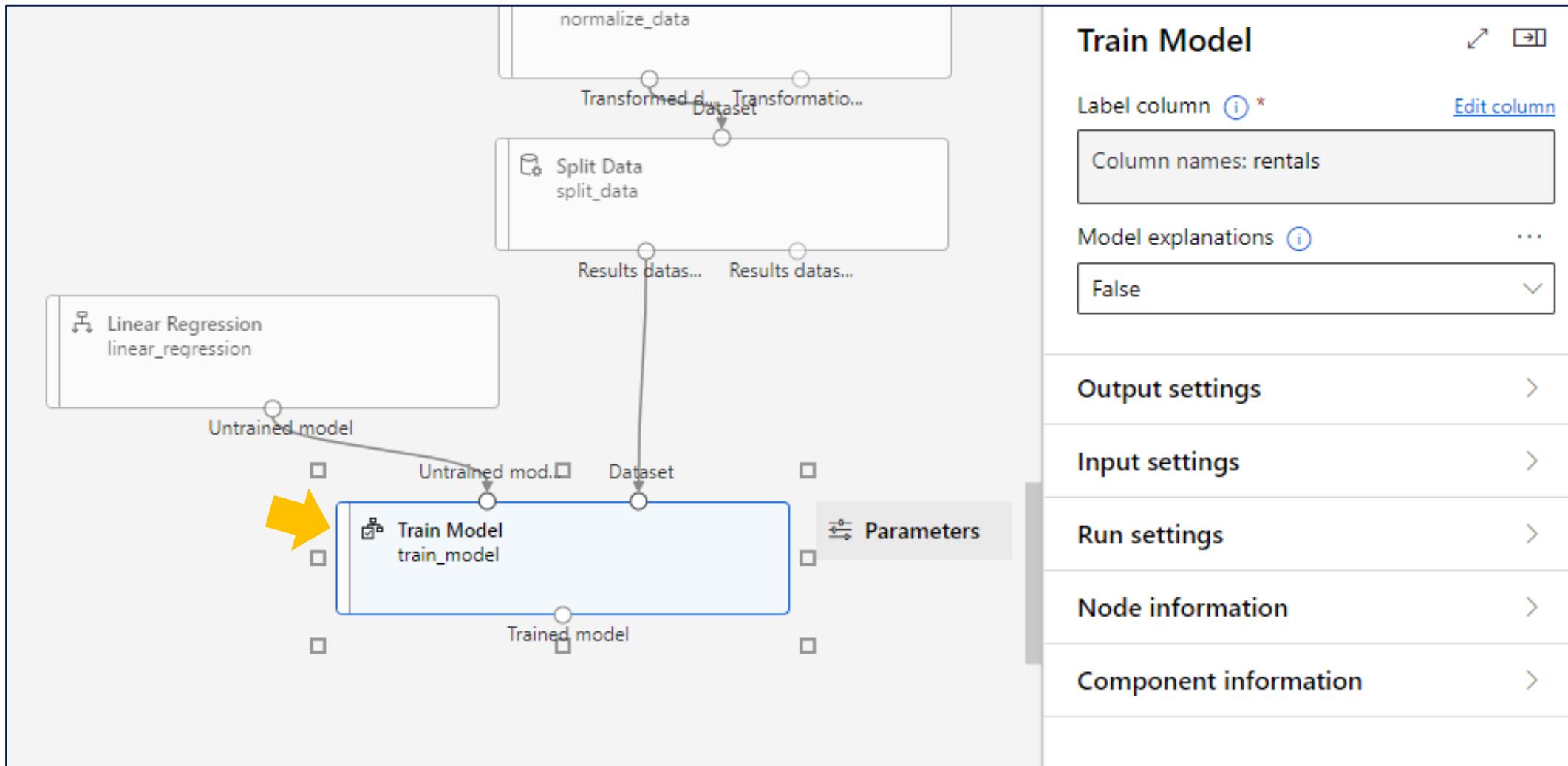
Name ↑	Category	Available quota (i)	Cost (i)
<input type="radio"/> Standard_E16ds_v4 16 cores, 128GB RAM, 600GB storage	Memory optimized	194 cores	\$1.38/hr
<input type="radio"/> Standard_E20ds_v4 20 cores, 160GB RAM, 750GB storage	Memory optimized	194 cores	\$1.73/hr
<input type="radio"/> Standard_E2ds_v4 2 cores, 16GB RAM, 75GB storage	Memory optimized	194 cores	\$0.17/hr
<input type="radio"/> Standard_E32ds_v4 32 cores, 256GB RAM, 1200GB storage	Memory optimized	194 cores	\$2.77/hr



Azure Machine Learning – Designer : 모델링



Azure Machine Learning – Designer : 모델링

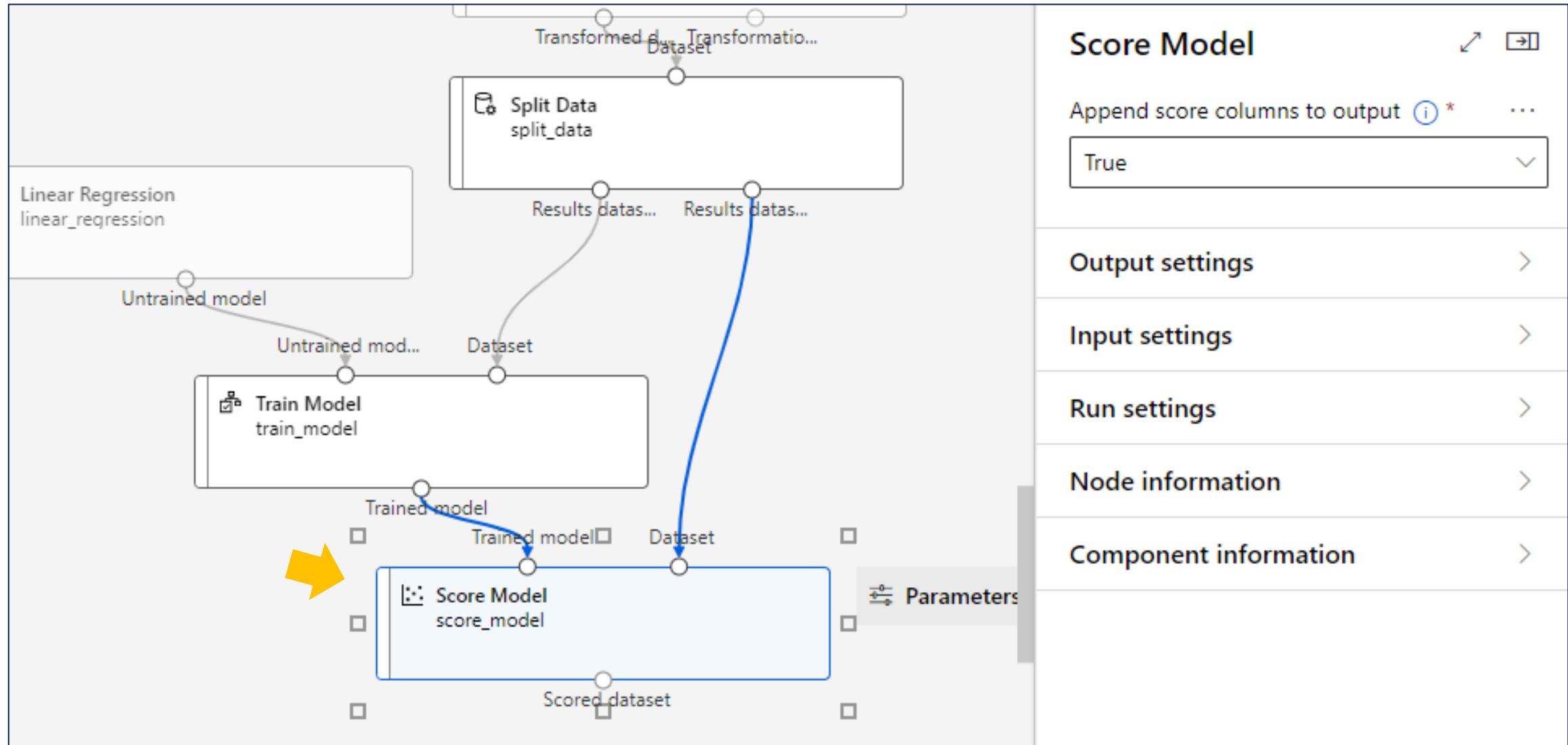


Azure Machine Learning – Designer : 머신러닝 절차

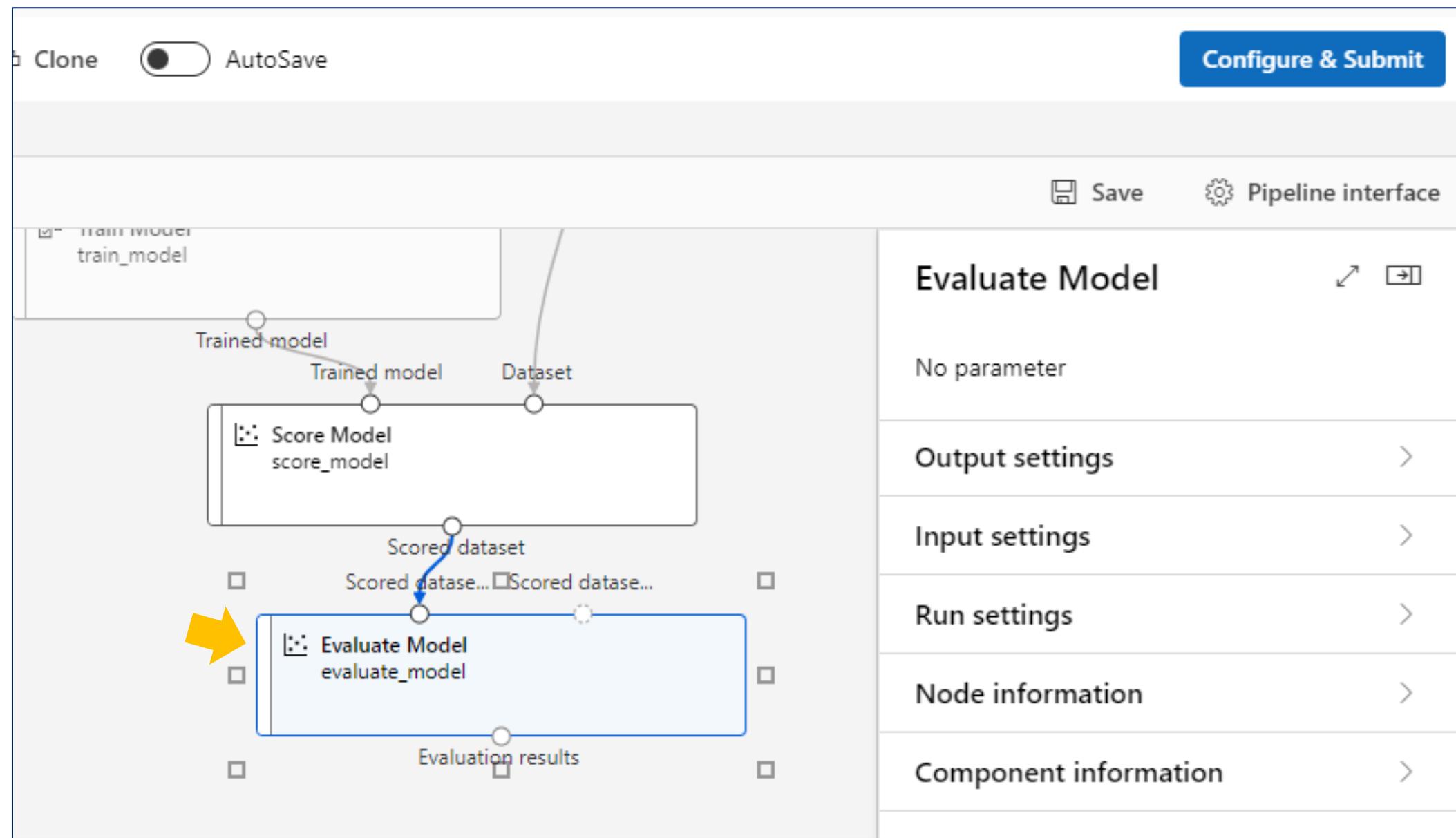


- | | | | | |
|--------------------------------------|---------------------------------------|-------------------------------------|--|----------------------|
| • 비즈니스
목적 정의 | • 데이터 수집
• 데이터 이해
• 데이터
레이블링 | • 데이터 전처리
• 데이터 분리
(학습 / 테스트) | • 모델링
알고리즘 선택
• 모델 학습(훈련)
• 모델 조정 | • 테스트
• 모델링 결과 평가 |
| • 환경/상황 분석 | | | | |
| • 모델링 방법 설정
(유형, 알고리즘
성능 등 검토) | | | | |

Azure Machine Learning – Designer : 평가



Azure Machine Learning – Designer : 평가



Azure Machine Learning – Designer : 평가

