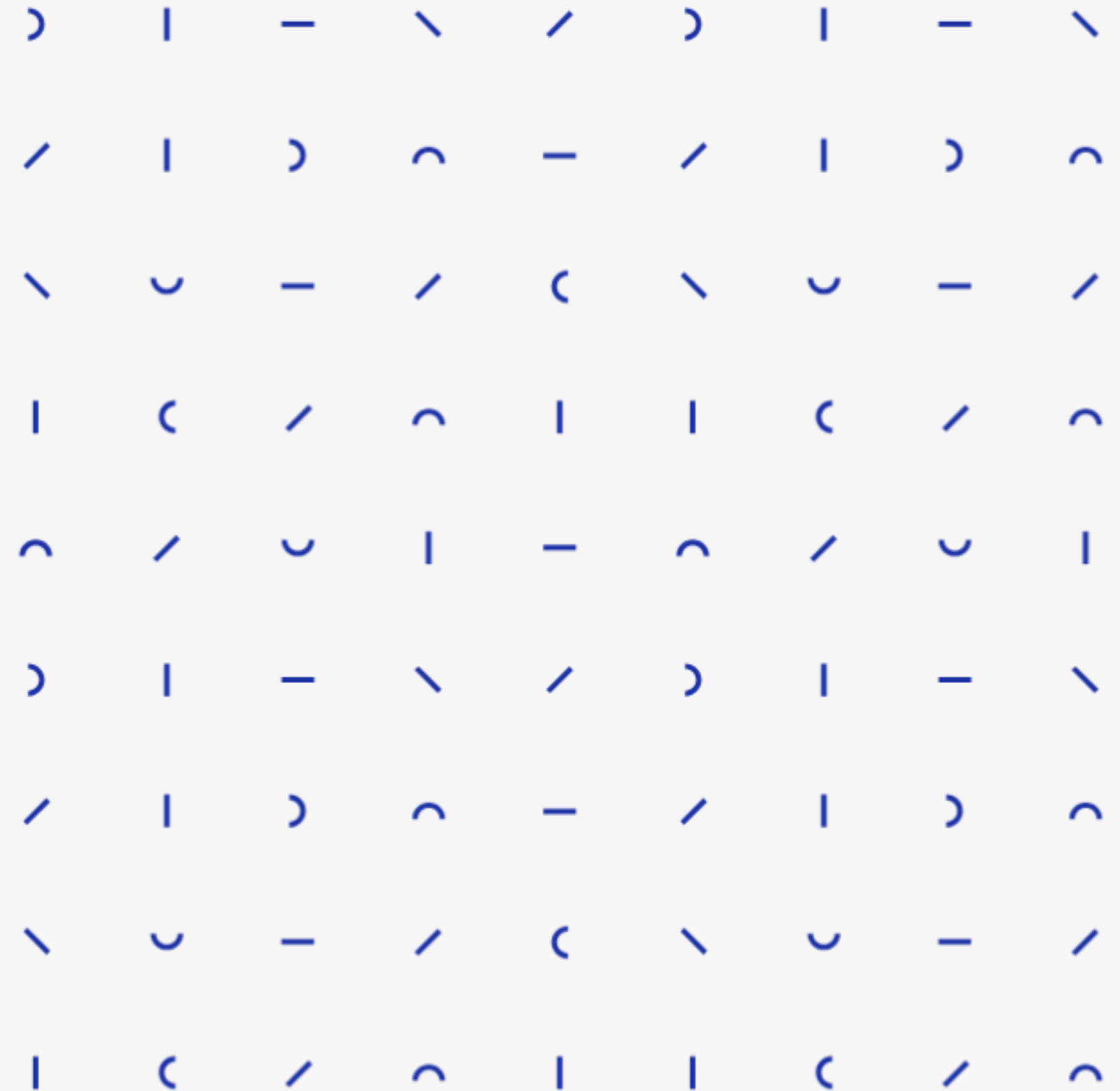


# Microsoft Azure

## Analytics & AI

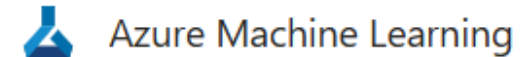
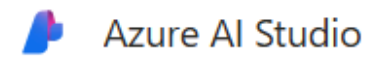


NGOC MY NGUYEN

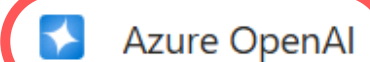
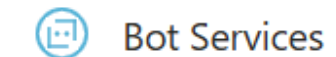
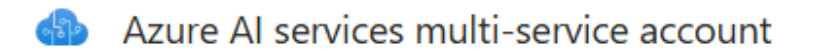
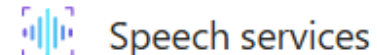
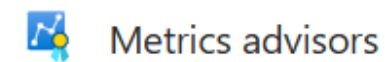
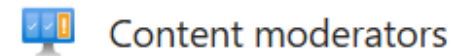
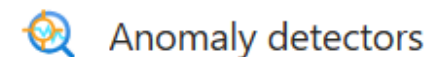
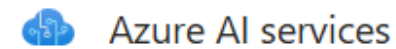
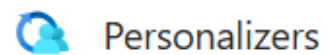
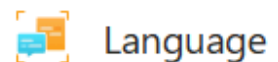
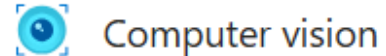
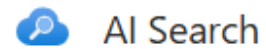


# AI + machine learning (22)

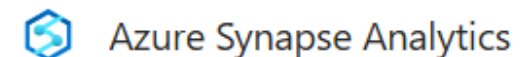
## Azure AI + Machine Learning platforms



## Azure AI services + APIs



## Industry Machine Learning solutions



# Azure OpenAI Services 🎂



## Funtionality

=> Build custom generative AI solutions with OpenAI's LLMs, image generation and embeddings model.



## Option 1: Standard (On-Demand) / Pay-as-you-go

=> Flexible, consumption-based pricing for variable workloads

(US/EU)	o1 Preview	o1 mini	GPT-4o	GPT-4o mini
Input	€15.7413	€3.1483	€2.62355	€0.15742
Cached input	€7.8707	€1.5742	€1.3118	€0.0792
Output	€62.965083	€12.593017	€10.4942	€0.6297

Table 1: Pricing per 1 Million tokens\* (as of Dec 2, 2024)

*\*token is a unit of text*

For English text:

- 1 token  
~ 4 characters  
~ 0.75 words
- 1 Million tokens  
~ 750,000 words  
~ 3000 pages of book  
(if 1 page = 250 words)  
~ all Harry Potter books

# Azure OpenAI Services 🎂



## Option 2: Provisioned Throughput Units (PTUs)

=> for workloads with consistent or predictable usage patterns  
=> providing stability and cost control

### Provisioned

You can allocate and manage throughput for deployments, ensuring predictable performance and stable capacity. You are charged an hourly rate per model regardless of usage, but you can also secure additional savings through monthly and annual reservations.

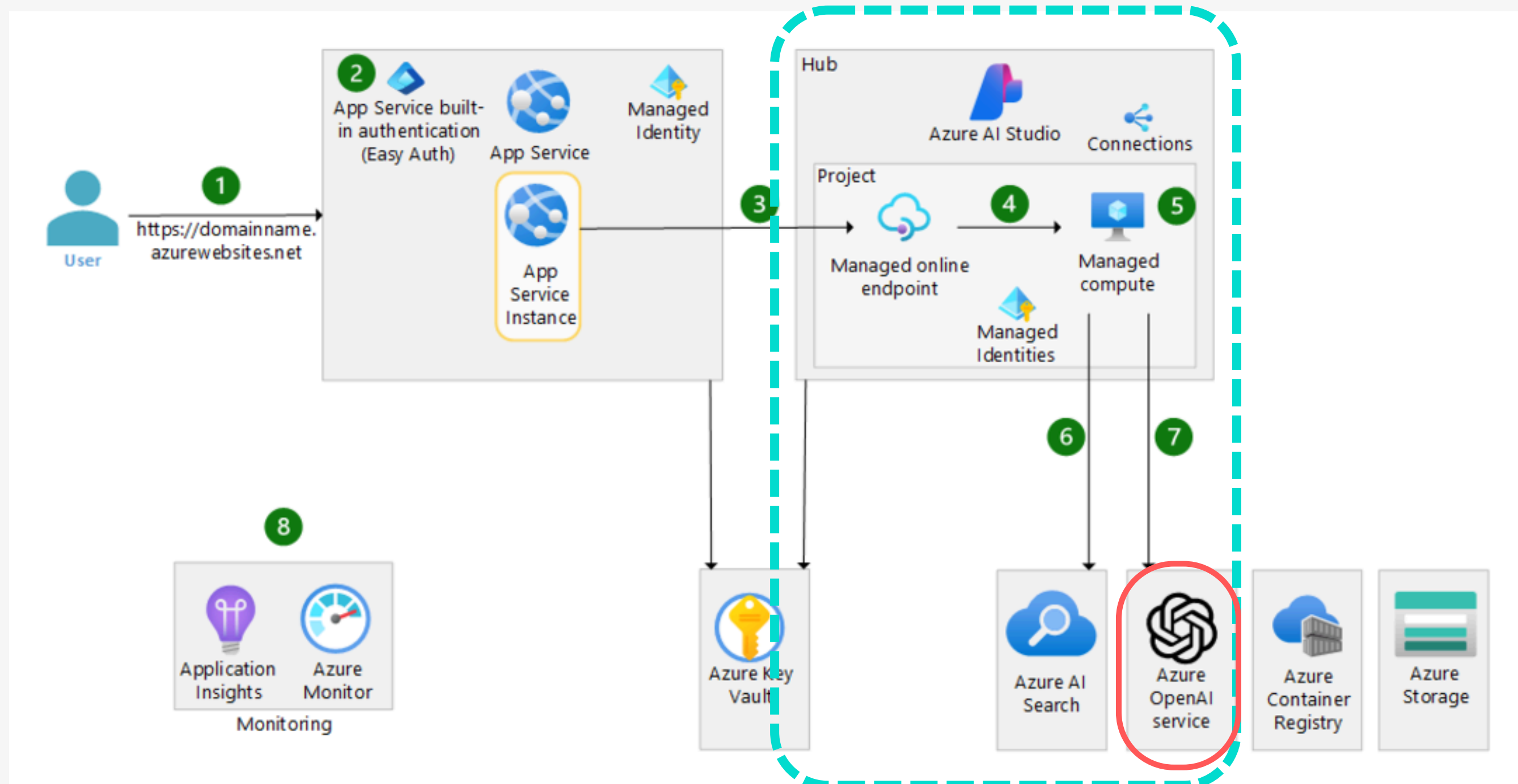
Model	Min PTUs	PTU Hourly pricing	PTU Monthly Reservation Pricing	PTU Yearly Reservation Pricing
GPT-4o Regional	50	€1.9081	€248.0443	€2,530.0516
GPT-4o Mini Regional	25	€1.9081	€248.0443	€2,530.0516

Table 2: Pricing via PTUs (as of Dec 2, 2024)



# Use Case example

Build a basic OpenAI end-to-end chat application (e.g. for your company)



# Live demo

## Chat playground

Help

[View code](#)

[Deploy](#)

[Import](#)

[Export](#)

[Prompt samples](#)

[Send feedback](#)

### Setup

Hide

#### Deployment \*

[+ Create new deployment](#)

gpt-4-32k (version:0613)

Give the model instructions and context ⓘ

You are an AI assistant that helps people find information.

[Apply changes](#)

[Generate prompt](#)

[+ Add section](#)

> Add your data

> Parameters

### Chat history



## Start with a sample prompt

#### 🔪 Recipe creation

Invent a recipe for a dish that combines flavors from two different cuisines.

#### 🕒 Historical fiction

Write a scene set in ancient Rome, focusing on the daily life of a common citizen.

Type user query here. (Shift + Enter for new line)

11/32768 tokens to be sent



Home

Get started

Model catalog

Playgrounds

Chat

Assistants PREVIEW

Real-time audio PREVIEW

Images

Completions

Tools

Fine-tuning

Azure OpenAI Evaluation PREVIEW

Batch jobs

Shared resources

Deployments

Quota

Safety + security











Data files

Vector



# Analytics (22)









## Big data processing

-  Analysis Services
-  Data Lake Analytics
-  HDInsight clusters
-  Microsoft Graph Data Connect PREVIEW
-  Apache Airflow™ on Astro - An Azure Native ISV ... PARTNER
-  Data Lake Storage Gen1
-  Azure HDInsight on AKS clusters (preview) PREVIEW
-  Data factories
-  Azure Databricks
-  Informatica Intelligent Data Management Cloud ... PARTNER

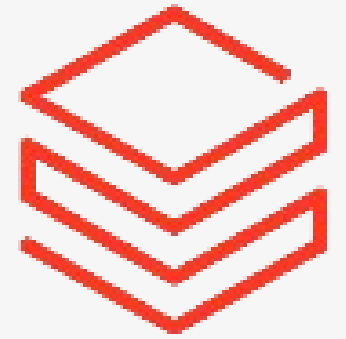
## Data exploration

-  Azure Data Explorer Clusters
-  Data Share Invitations
-  Data Shares
-  Power BI Embedded

## Real-time analytics

-  Apache Kafka® & Apache Flink® on Confluent ... PARTNER
-  Event Hubs
-  Log Analytics workspaces
-  Managed Prometheus
-  Stream Analytics clusters
-  Stream Analytics jobs
-  Azure Synapse Analytics
-  Azure Synapse Analytics (private link hubs)

# Azure Databricks



## **Functionality**

=> A unified, open analytics platform for building, deploying, sharing, and maintaining enterprise-grade data, analytics, and AI solutions at scale.



**ETL & Data Engineering**



**AI and Machine Learning**



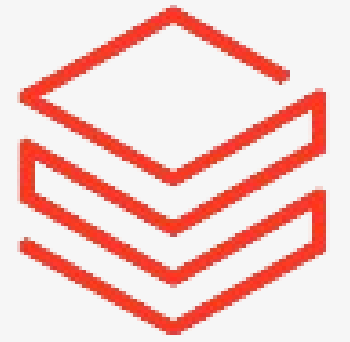
**Data Warehousing and Analytics**



**Dashboard/Visualization**



# Azure Databricks



charges for the VMs (virtual machines) provisioned + the DBUs\* (Databricks Units)  
=> Prices differ based on:

## Type of tier:

- **Standard** tier

=> small team, running non-critical workloads, for experimenting

- **Premium** tier

=> enhanced security, governance, for operations in a regulated environment requiring compliance.

## Type of workload:

- Interactive Serverless Compute
- All-Purpose Compute (with Photon)
- Automated Serverless Compute
- Jobs Compute (with Photon)
- Jobs Light Compute
- SQL Compute
- SQL Pro Compute
- Serverless SQL
- Serverless Real-Time Inference
- Model Training
- Databricks Storage Unit
- Delta Live Tables (DLTs)

## VM instance:

- DSv2 series: DS3 v2, DS4 v2, DS5 v2
- Dv2 series: D3 v2, D4 v2, D5 v2
- Dsv3 series: D4s v3, D8s v3, D16s v3,  
...
- .....


varied by vCPU and RAM

## & Region

*\*DBU = unit of processing capability,  
measure of compute power*

# Live demo


Microsoft Azure


 databricks

Search data, notebooks, recents, and more...

CTRL + P

nmn-databricks



 N

New

Workspace

Recents

Catalog

Workflows

Compute

SQL

SQL Editor

Queries

Dashboards

Genie

Alerts

Query History

SQL Warehouses

Data Engineering

Job Runs

Data Ingestion

Delta Live Tables

Machine Learning

Playground

Experiments

Features

Catalog

Type to search...

For youAll

My organization

- nmn\_databricks
- system

Shared

- samples

Legacy

- hive\_metastore

nmn-demo1nmn-demo2

Run all (1000)

nmn\_dat...def...

Serverless Starter...ServerlessS

SaveScheduleShare

```
1 USE CATALOG samples;
2 SELECT UPPER(country_code) AS country_code, temperature_avg
3 FROM samples.accuweather.forecast_daily_calendar_metric
4 WHERE date = '2024-07-13';
```

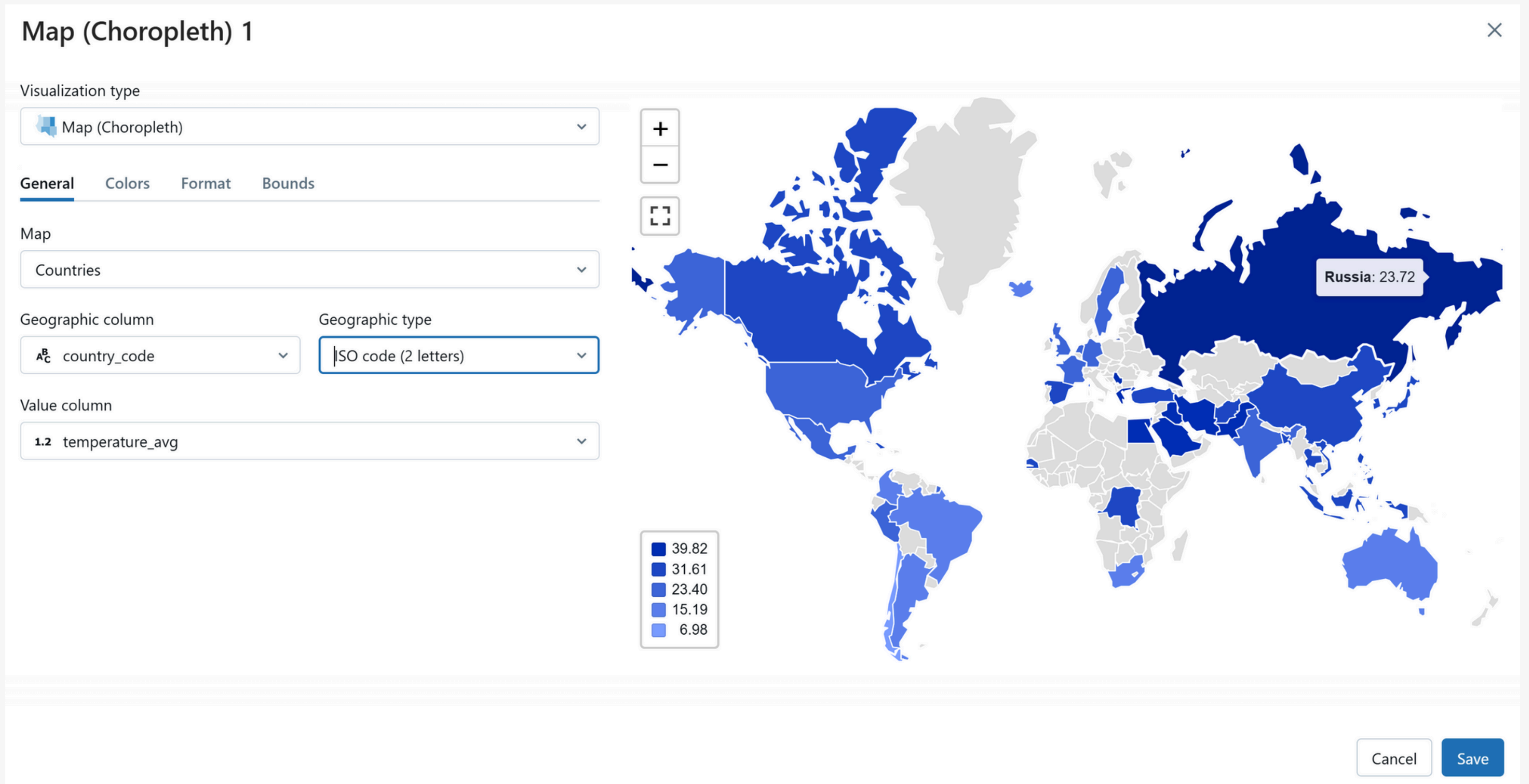
Raw results

	country_code	temperature_avg
1	IN	30.66
2	ZA	11.39
3	JP	26.82
4	TR	28.42
5	CL	6.98
6	PE	15.81
7	IR	31.82
8	TH	29.53

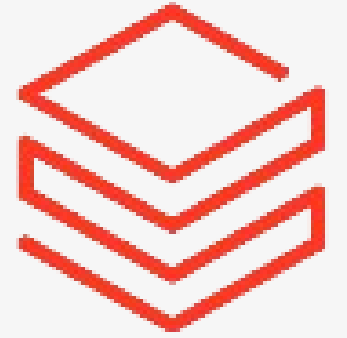
4 s 892 ms | 50 rows returned

Refreshed 16 hours ago

# Live demo



# Azure Databricks



Run **Premium** tier cluster  
for **100 hours**  
in **East US 2** with **10 DS13v2** instances

## Pay as You Go

**All-Purpose Compute** workload:

- VM cost for 10 DS13v2 instances

*100 hours x 10 instances x €0.571/hour = €570.502*

- DBU cost for All-Purpose Compute workload  
for 10 DS13v2 instances

*100 hours x 10 instances x 2.0 DBU per instance per  
hour x €0.525/DBU = €1,049.419*

- **Total cost:** €570.502 (VM Cost) + €1,049.419  
(DBU Cost) = **€1,619.920.**

(as of Dec 2, 2024)



## Option 1: Pay-as-you-go

=> Flexible, consumption-  
based pricing for variable  
workloads

## Option 2: Pre-purchase plan

=> 1 or 3 years

=> with discounts

=> Certain number of DBCUs\*  
(Databricks Commit Unit)

*\*DBCU = prepaid unit of compute,  
essentially a billing credit.*

*1 DBCU = the cost of running one DBU*



# Thank *you*!

Feel free to ask questions if you have any.

