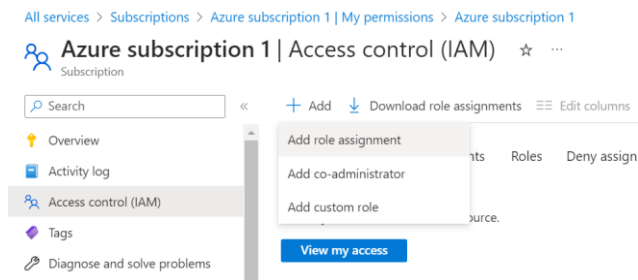


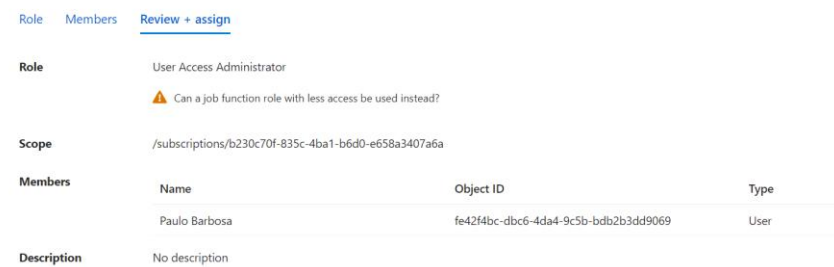
Step by step to run the application

1. Open Azure Portal
2. Go to Access Control (IAM) in your Subscription options
3. Click on Add Role Assignment



4. Add “User Access Administrator”, “Storage Blob Data Contributor” and “Key Vault Administrator” Roles

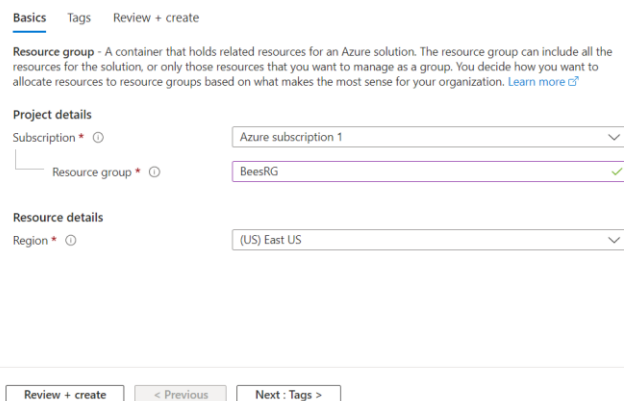
Add role assignment



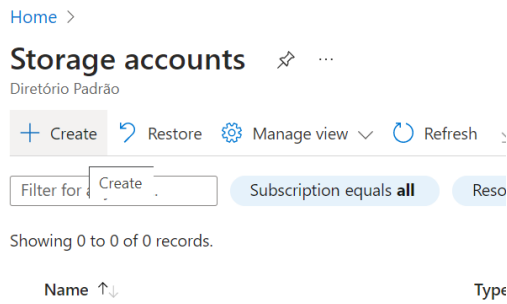
5. Go to “Resource Groups” and create a resource group

Home > Subscriptions > Azure subscription 1 | Resource groups >

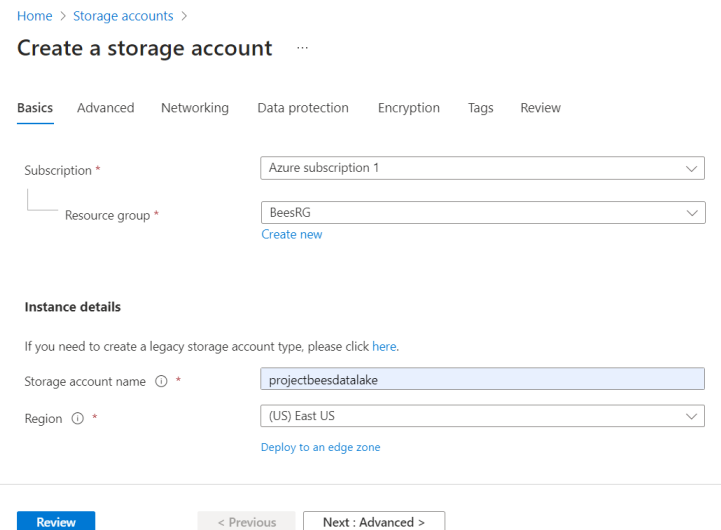
Create a resource group



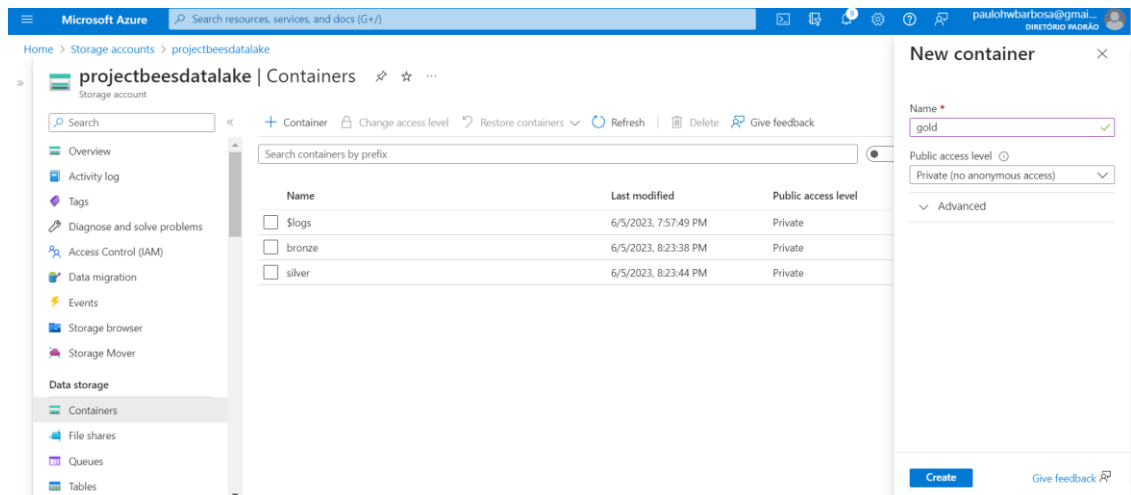
6. Go to Storage accounts and click on “create”



7. Create the Storage Account as the image below, the Storage account name must be unique:



8. Access this Storage Account, go to “Containers” and create a bronze, a silver and a gold container



9. Go to “Access Keys” and copy the Key value to a notepad.

Home > projectbeesdatalake_1686005826078 | Overview > projectbeesdatalake

projectbeesdatalake | Access keys

Storage account

Search

Storage browser
Storage Mover

Data storage

Containers
File shares
Queues
Tables

Security + networking

Networking
Azure CDN
Access keys
Shared access signature

Set rotation reminder Refresh Give feedback

Access keys authenticate your applications' requests to this storage account. Keep your keys in a secure location like Azure Key Vault, and replace them often with new keys. The two keys allow you to replace one while still using the other.

Remember to update the keys with any Azure resources and apps that use this storage account.
[Learn more about managing storage account access keys](#)

Storage account name
projectbeesdatalake

key1 Rotate key
Last rotated: 05/06/2023 (0 days ago)
Key
PDObifir8h5GBxEnjJ8eDcsL2bsioK/alM8L+A30t3kD/70GSbzS0cMu8Nz57QvMLV... Hide
Connection string
DefaultEndpointsProtocol=https;AccountName=projectbeesdatalake;AccountKey=... Hide

key2 Rotate key

10. Go to Data protection and disable Soft Delete for blobs

Home > Storage accounts > projectbeesdatalake

projectbeesdatalake | Data protection

Storage account

Search

Access keys
Shared access signature
Encryption
Microsoft Defender for Cloud

Data management

Redundancy
Data protection
Object replication
Blob inventory
Static website
Lifecycle management
Azure search

Settings

Data protection provides options for recovering your data when it is erroneously modified or deleted.

Recovery

☐ Enable operational backup with Azure Backup

☐ Enable point-in-time restore for containers

☐ Enable soft delete for blobs
Soft delete enables you to recover blobs that were previously marked for deletion, including blobs that were overwritten. [Learn more](#)
You will still be able to access and recover soft deleted data if the soft delete feature is disabled.

☒ Enable soft delete for containers
Soft delete enables you to recover containers that were previously marked for deletion. [Learn more](#)
Keep deleted containers for (in days) 7

☐ Enable permanent delete for soft deleted items

Save Discard

11. Go to Azure Active Directory, App registrations and "New registration" to create a service principal

Home > Diretório Padrão

Diretório Padrão | App registrations

Azure Active Directory

Manage

Users
Groups
External Identities
Roles and administrators
Administrative units
Delegated admin partners
Enterprise applications
Devices
App registrations
Identity Governance
Application proxy

+ New registration Endpoints Troubleshooting Refresh Download Preview features Got feedback?

New registration

Starting June 30th, 2020 we will no longer add any new features to Azure Active Directory Authentication Library (ADAL) and Azure AD Graph. We security updates but we will no longer provide feature updates. Applications will need to be upgraded to Microsoft Authentication Library (MSAL)

All applications **Owned applications** Deleted applications Applications from personal account

Start typing a display name or application (client) ID to filter these r... Add filters

This account isn't listed as an owner of any applications in this directory.

[View all applications in the directory](#)
[View all applications from personal account](#)

12. Fill the gaps as the image below:

[Home](#) > [Diretório Padrão](#) | [App registrations](#) >

Register an application ...

bees-sp ✓

Supported account types

Who can use this application or access this API?

- ☒ Accounts in this organizational directory only (Diretório Padrão only - Single tenant)
- ☐ Accounts in any organizational directory (Any Azure AD directory - Multitenant)
- ☐ Accounts in any organizational directory (Any Azure AD directory - Multitenant) and personal Microsoft accounts (e.g. Skype, Xbox)
- ☐ Personal Microsoft accounts only

[Help me choose...](#)

Redirect URI (optional)

We'll return the authentication response to this URI after successfully authenticating the user. Providing this now is optional and it can be changed later, but a value is required for most authentication scenarios.

Web ▼ e.g. https://example.com/auth ✓

Register an app you're working on here. Integrate gallery apps and other apps from outside your organization by adding from [Enterprise applications](#).

By proceeding, you agree to the [Microsoft Platform Policies](#) ↗

Register

13. Go to "Certificates & secrets" and create a client secret, and copy those values to a notepad

[Home](#) > [Diretório Padrão](#) | [App registrations](#) > bees-sp

bees-sp | Certificates & secrets ✕ ...

Search

«

Got feedback?

Overview

Quickstart

Integration assistant

Manage

Branding & properties

Authentication

Certificates & secrets

Token configuration

API permissions

Expose an API

App roles

Owners

Roles and administrators

Got a second to give us some feedback? →

Credentials enable confidential applications to identify themselves to the authentication service when receiving tokens at a web addressable location (using an HTTPS scheme). For a higher level of assurance, we recommend using a certificate (instead of a client secret) as a credential.

Application registration certificates, secrets and federated credentials can be found in the tabs below.

Certificates (0) **Client secrets (1)** Federated credentials (0)

A secret string that the application uses to prove its identity when requesting a token. Also can be referred to as application password.




+ New client secret

Description	Expires	Value	Secret ID
bees-sp	9/3/2023	XE28Q~GzG2mMGAM-QYpqByrCFBXjV3...	6417e00a-57ec-4160-bceb-4757f7691aee

14. Go to Overview and copy the Application (client) ID and the Directory (tenant) ID to a notepad

Home > Diretório Padrão | App registrations >

 bees-sp ✎ ...

«  Delete  Endpoints  Preview features

Overview

Quickstart

Integration assistant

Manage

Branding & properties

Authentication

...

^ Essentials

Display name : [bees-sp](#)

Application (client) ID : 1729b148-7b96-4189-85e9-c038dedba45d

Object ID : 99c3e806-945b-4e5b-a9f6-e3a521f6a Copied

Directory (tenant) ID : 8bd0fed4-7de6-4dc1-ae38-0bf1d336da5a

Supported account types : [My organization only](#)

15. Go to Key Vaults and create a key vault

Home > Key vaults >

Create a key vault

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription *

Azure subscription 1

Resource group *

BeesRG

Create new

Instance details

Key vault name * ⓘ

BeesProjectKV ✓

Region *

East US




Pricing tier * ⓘ

Standard

16. Access the created Key Vault and click on “Generate/Import”, in the Secrets section

Home > BeesKV

 **BeesKV | Secrets** ☆ ...
Key vault

«  Generate/Import  Refresh 

Events

Objects

Keys

Secrets

Certificates

Name

There are no secrets available.

17. We are going to create some secrets:

Secret adb-tenantId using Directory (tenant) Id value copied at step 14.

Secret adb-appld using Application (client) Id value copied at step 14.

Secret adb-appSecret using App's Secret value copied at step 13.

[Home](#) > [BeesProjectKV | Secrets](#) >

Create a secret ...

Upload options	<div>Manual</div>
Name * ⓘ	<div>adb-appKey</div> <div>✖ The name 'adb-appKey' is already in use.</div>
Secret value * ⓘ	<div>.....</div>
Content type (optional)	<div></div>
Set activation date ⓘ	<div><input type="checkbox"/></div>
Set expiration date ⓘ	<div><input type="checkbox"/></div>
Enabled	<div>Yes No</div>
Tags	<div>0 tags</div>

18. Go to Access configuration and select “Vault access policy”

Home > [BeesProjectKV](#)

BeesProjectKV | Access configuration

☆ ...

Key vault

Search

<<

Refresh

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Access policies

Events

Objects

Keys

Secrets

Certificates

Settings

Access configuration

Configure your options on access policy for this key vault

To access a key vault in data plane, all callers (users or applications) must have proper authn can execute. [Learn more](#)

Permission model

Grant data plane access by using a [Azure RBAC](#) or [Key Vault access policy](#)

☐ Azure role-based access control (recommended) ⓘ

☒ Vault access policy ⓘ

Go to access policies

Resource access

Choose among the following options to grant access to specific resource types

☐ Azure Virtual Machines for deployment ⓘ

19. Go to Access policies and click on create

Home > BeesProjectKV

BeesProjectKV | Access policies

Key vault

[+ Create](#) [Refresh](#)

- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Access policies**
- Events

Access policies enable you to ha

Showing 1 to 2 of 2 records.

☐ Name ↑↓

☒ APPLICATION

☐ AzureDatabricks

20. Select Key & Secret management to the service principal

Home > BeesProjectKV | Access policies >

Create an access policy

BeesProjectKV

Configure from a template

Key & Secret Management

Key permissions

Key Management Operations

☒ Select all

- ☒ Get
- ☒ List
- ☒ Update
- ☒ Create
- ☒ Import
- ☒ Delete
- ☒ Recover
- ☒ Backup
- ☒ Restore

Cryptographic Operations

Secret permissions

Secret Management Operations

☒ Select all

- ☒ Get
- ☒ List
- ☒ Set
- ☒ Delete
- ☒ Recover
- ☒ Backup
- ☒ Restore

Privileged Secret Operations

- ☐ Select all
- ☐ Purge

Certificate permissions

Certificate Management Operations

☐ Select all

- ☐ Get
- ☐ List
- ☐ Update
- ☐ Create
- ☐ Import
- ☐ Delete
- ☐ Recover
- ☐ Backup
- ☐ Restore
- ☐ Manage Contacts
- ☐ Manage Certificate Authorities

[Home](#) > [BeesProjectKV | Access policies](#) >

Create an access policy ...

BeesProjectKV

✓ Permissions **2 Principal** ③ Application (optional) ④ Review + create

Only 1 principal can be assigned per access policy.

Use the new embedded experience to select a principal. The previous popup experience can be accessed here. [Select a principal](#)

 bees-sp
1729b148-7b96-4189-85e9-c038dedba45d


 beesADF
c58a306a-d11f-4bcb-bd4f-d3b24e8e9e15

Selected item

 bees-sp
1729b148-7b96-4189-85e9-c038dedba45d

21. Go to the Storage account and go to Access Control (IAM) and click on “Add”

[Home](#) > [projectbeesdatalake](#)

 **projectbeesdatalake** | Access Control (IAM) ☆

Storage account

[+ Add](#) [Download role assignments](#)

[Overview](#) [Activity log](#) [Tags](#) [Diagnose and solve problems](#) **[Access Control \(IAM\)](#)** [Data migration](#) [Export](#)

Check access Role assignments Roles

My access
View my level of access to this resource.
[View my access](#)

Check access
Review the level of access a user, group, service pri

22. Add Storage Blob Data Contributor and Key Vault Administrator roles to the service principal

[Home](#) > [projectbeesdatalake | Access Control \(IAM\)](#) >

Add role assignment ...

Role **Members** Conditions (optional) Review + assign

Selected role Storage Blob Data Contributor

Assign access to ☒ User, group, or service principal
☐ Managed identity

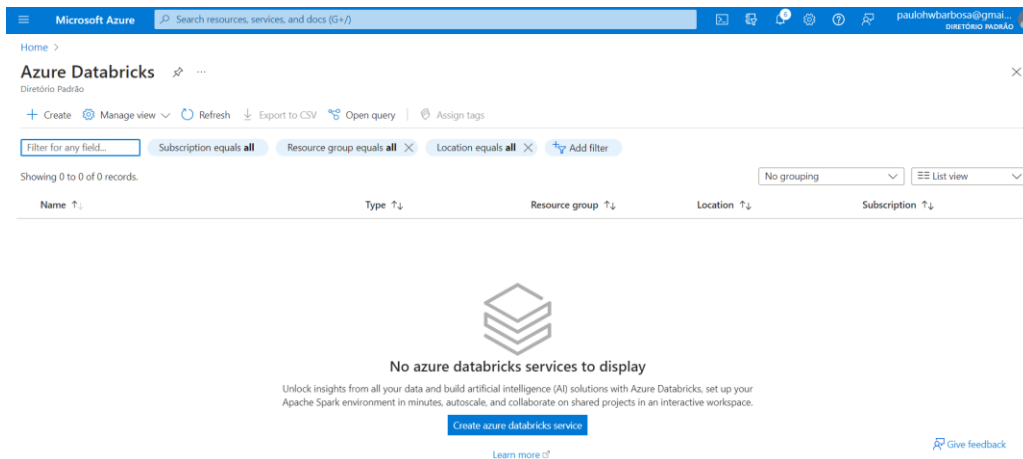
Members [+ Select members](#)

Name	Object ID	Type
bees-sp	1d2dbfae-9123-4edb-9769-3983e7a1e1...	App

Description

Optional

23. Go to Azure Databricks and click on create Azure Databricks service



24. Create the Databricks Workspace as the image below:

[Home](#) > [Azure Databricks](#) >

Create an Azure Databricks workspace

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * ⓘ

Resource group * ⓘ [Create new](#)

Instance Details

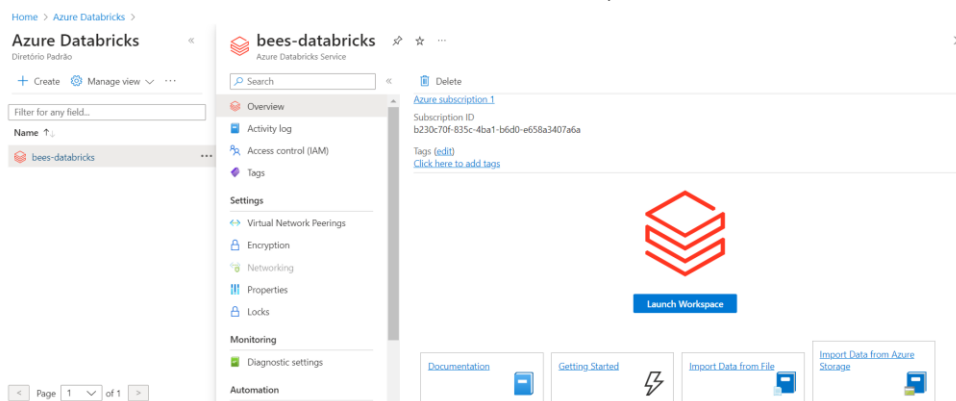
Workspace name * ✓

Region *

Pricing Tier * ⓘ

i We selected the recommended pricing tier for your workspace. You can change the tier based on your needs. **x**

25. Go to Azure Databricks and click on “Launch Workspace”



26. Go to the following URL and fill de gaps as shown on the image bellow:

<https://<databricks-instance>#secrets/createScope>

Microsoft Azure | databricks | Search data, notebooks, recents, and more...

HomePage / Create Secret Scope

Create Secret Scope | Cancel | Verifying...

A store for secrets that is identified by a name and backed by a specific store type. [Learn more](#)

Scope Name [?]

akv-scope

Manage Principal [?]

All Users

Azure Key Vault [?]

DNS Name

https://bees.vault.azure.net/

Resource ID

/subscriptions/4fe0fff3-5fb2-4514-87ea-522c8566beaa/resourceGroups/default/pro

DNS Name field fill with Vault URI value and Resource ID with Resource ID Value, found in the Properties section of the created Key Vault

Microsoft Azure | Search resources, services, and docs (G+/I)

Home > Key vaults > BeesKV

Key vaults

BeesKV

BeesKV | Properties

Search

Save | Discard changes | Refresh

Certificates

Settings

- Access configuration
- Networking
- Microsoft Defender for Cloud
- Properties
- Locks

Monitoring

- Alerts
- Metrics
- Diagnostic settings
- Logs
- Insights
- Workbooks

Name

BeesKV

Sku (Pricing tier)

Standard

Location

brazilsouth

Vault URI

https://beeskv.vault.azure.net/

Resource ID

/subscriptions/b230c70f-835c-4ba1-b6d0-e658a3407a6a/resourceGroups/BeesRG/pr...

Subscription ID

b230c70f-835c-4ba1-b6d0-e658a3407a6a

Subscription Name

Azure subscription 1

Directory ID

6bd0fed4-7de6-4dc1-ae38-0bf1d336da5a

Directory Name

Diretório Padrão

Soft-delete

Soft delete has been enabled on this key vault

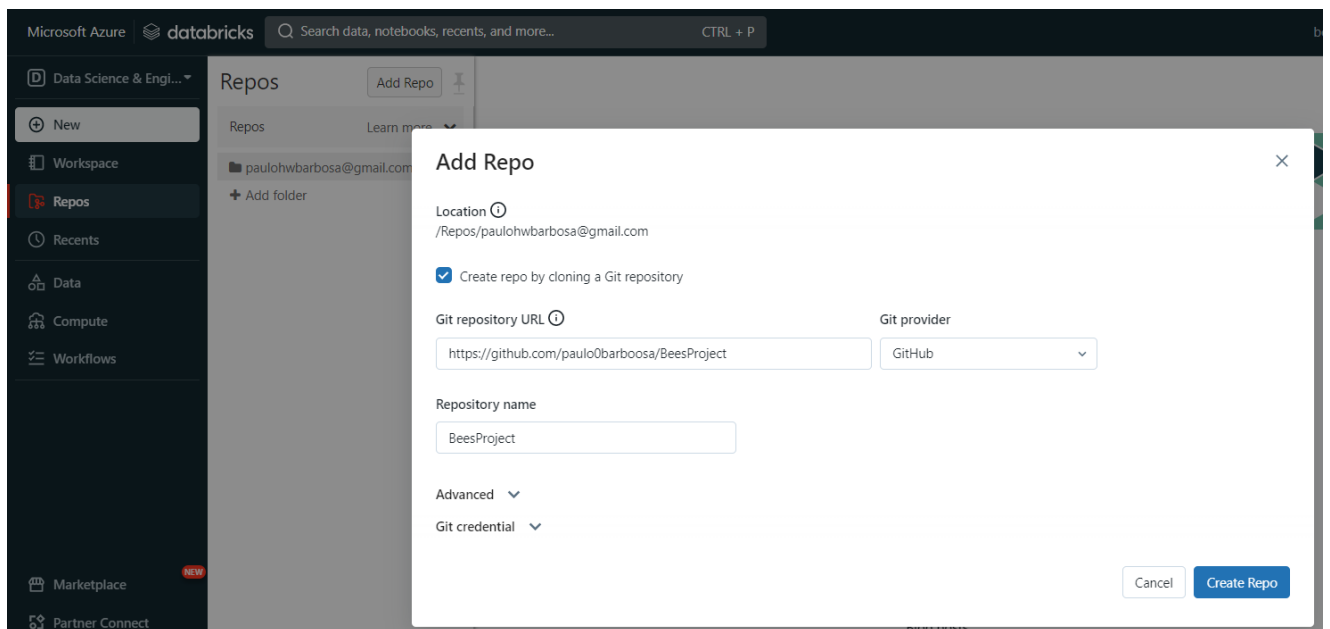
Days to retain deleted vaults

90

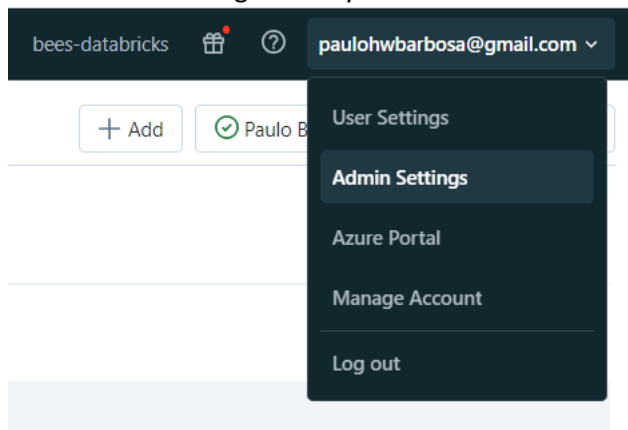
Purge protection

Disable purge protection (allow key vault and objects to be purged during retention)

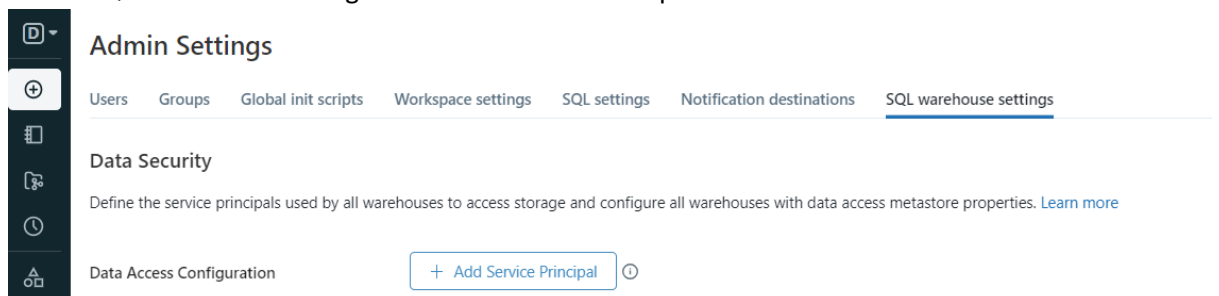
27. Go to “Repos” Section of Databricks, select “Add Repo” and fill the fields as the image below:



28. Go to Admin Settings under your name account



29. Go to SQL warehouse settings and add a Service Principal



30. Fill the gaps as the image below (change Storage Account Name to your Storage Account):

Add Service Principal
X

Storage Account Name ⓘ

Application (client) ID ⓘ

Directory (tenant) ID ⓘ

Application Client Secrets

Secret Scope ⓘ

[Create a secret scope](#)

Secret Key ⓘ

[Create a secret for your application client secret](#)

Cancel
Add

31. Inside both silver and gold notebooks, change the first variable to your Storage Account Name.

Data Science & Engi...
New
Workspace
Repos
Recents
Data
Compute
Workflows

Repos

BeesProject main

databricks

README.md

databricks

gold

silver

Add Repo

Cmd 1

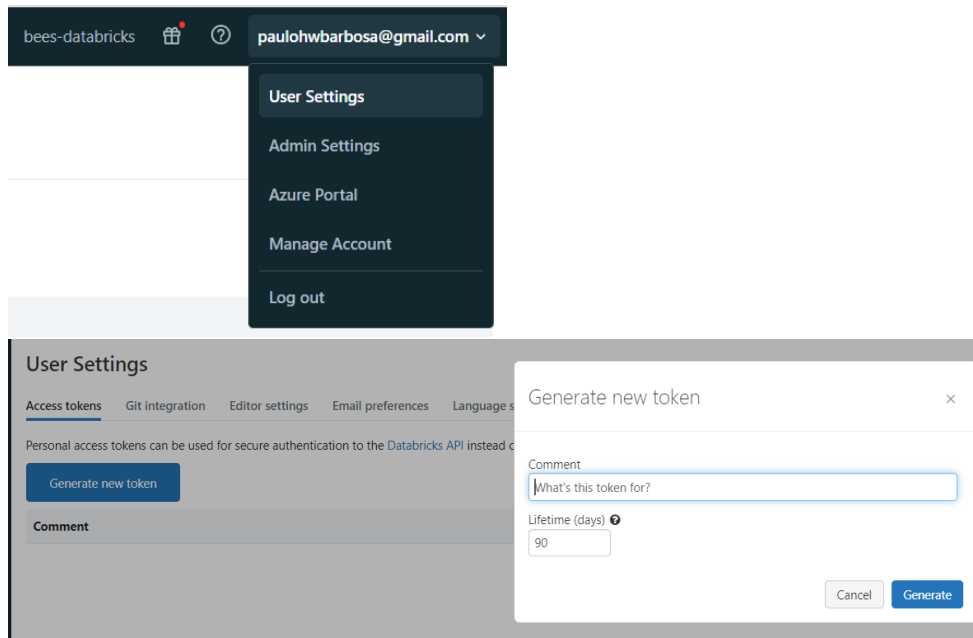
Defines the Storage Account Name, Secrets and Spark Configurations

```

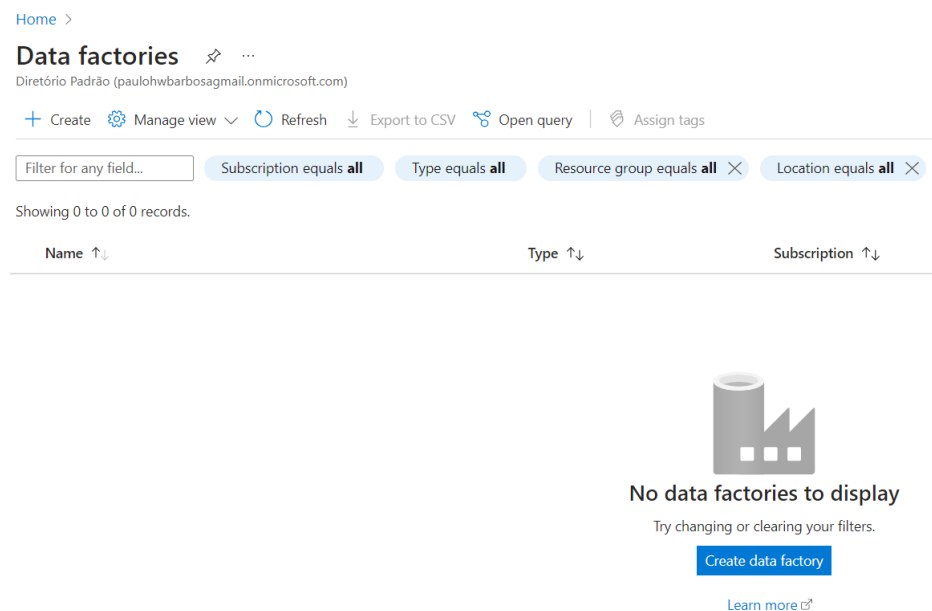
1 storage_account_name = "projectbeesdatalake"
2 appid = dbutils.secrets.get(scope = "sp-scope", key = "adb-appId") #App ID
3 appsecret = dbutils.secrets.get(scope = "sp-scope", key = "adb-appSecret") #App Secret
4 tenantid = dbutils.secrets.get(scope = "sp-scope", key = "adb-tenantId") #Tenant ID
5

```

32. Go to “User Settings” and generate an access token, copy it and save it on a notepad



33. Go to “Data Factories” in the Azure Portal and “Create data factory”



34. Create as the images below:

[Home](#) > [Data factories](#) >

Create Data Factory

Basics Git configuration Networking Advanced Tags Review + create

One-click to create data factory with sample pipeline and datasets. [Try it](#)

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * ⓘ	<div>Azure subscription 1</div>
Resource group * ⓘ	<div>BeesRG</div> <div>Create new</div>

Instance details

Name * ⓘ	<div>beesADF</div>
Region * ⓘ	<div>East US</div>
Version * ⓘ	<div>V2</div>

Basics **Git configuration** Networking Advanced Tags Review + create

Azure Data Factory allows you to configure a Git repository with either Azure DevOps or GitHub. Git is a version control system that allows for easier change tracking and collaboration.

[Learn more about Git integration in Azure Data Factory](#)

Configure Git later ⓘ	<input type="checkbox"/>
Repository Type * ⓘ	<div><input type="radio"/> Azure DevOps</div> <div><input checked="" type="radio"/> GitHub</div>
GitHub account * ⓘ	<div>paulo0barboosa</div>
Repo name * ⓘ	<div>BeesProject</div>
Branch name * ⓘ	<div>main</div>
Root folder * ⓘ	<div>/</div>

[Review + create](#)

[< Previous](#)

[Next : Networking >](#)

35. Launch your Azure Data Factory

Microsoft Azure Upgrade Search resources, services, and docs (G+)

Home > Microsoft.DataFactory-20230605222328 | Overview >

beesADF

Data factory (V2)

Search

Delete

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Settings

Networking

Managed identities

Properties

Locks

Getting started

Quick start

Monitoring

Essentials

Resource group (move) : BeesRG

Status : Succeeded


Location : East US

Subscription (move) : Azure subscription 1

Subscription ID : b230c70f-835c-4ba1-b6d0-e658a3407a6a

Type : Data factory (V2)

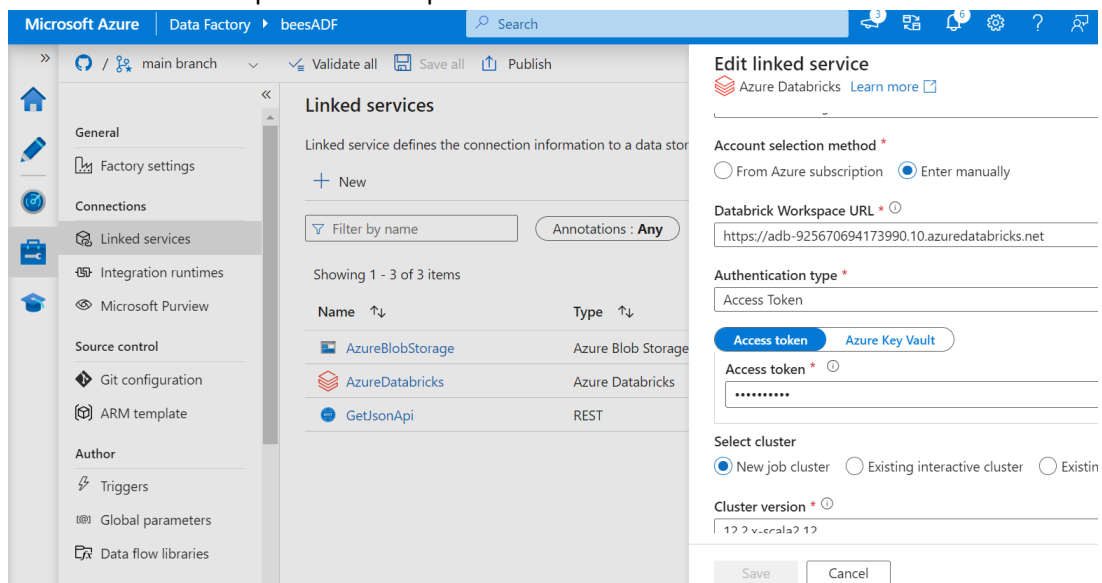
Getting started : [Quick start](#)



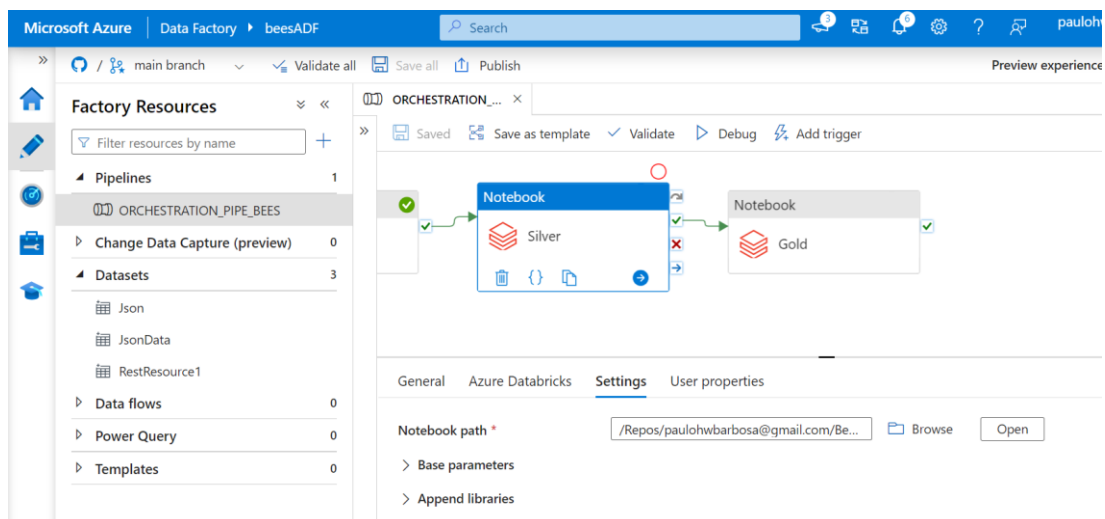
Azure Data Factory Studio

Launch studio

36. In the Manage options, select Linked Services, AzureDatabricks and change the Access Token value to the value copied at the step 32



37. In the Author options, select Pipelines, ORCHESTRATION_PIPE_BEES and change the notebook path at the settings of Silver and Gold Databricks Jobs to your email account
Silver Notebook: /Repos/<account>/BeesProject/databricks/silver
Gold Notebook: /Repos/<account>/BeesProject/databricks/gold



38. Now you can click on “Debug” to run the pipeline

Microsoft Azure

Data Factory

beesADF

Search

main branch

Validate all

Save all

Publish

Preview experience

C

Factory Resources

Filter resources by name

Pipelines1

ORCHESTRATION_PIPE_BEES

Change Data Capture (preview)0

Datasets3

Json

JsonData

RestResource1

Data flows0

Power Query0

Templates0

ORCHESTRATION_PIPE_BEES

Saved

Save as template

Validate

Debug

Add trigger

Copy data

Copy data

Notebook

Notebook

RawData

Bronze

Silver

Gold

Parameters

Variables

Settings

Output

Pipeline run ID: d96b86b1-867e-4c34-8993-37f9deb5ebda

View debug run console

All status

Export to CSV

Showing 1 - 2 of 2 items

Activity name	Status	Activity type	Run start	Duration