**Azure Data Engineering Project**

The tools that are covered in this project are,

1. Azure Data Factory

2. Azure Data Lake Storage Gen2

3. Azure Databricks

4. Azure Synapse Analytics

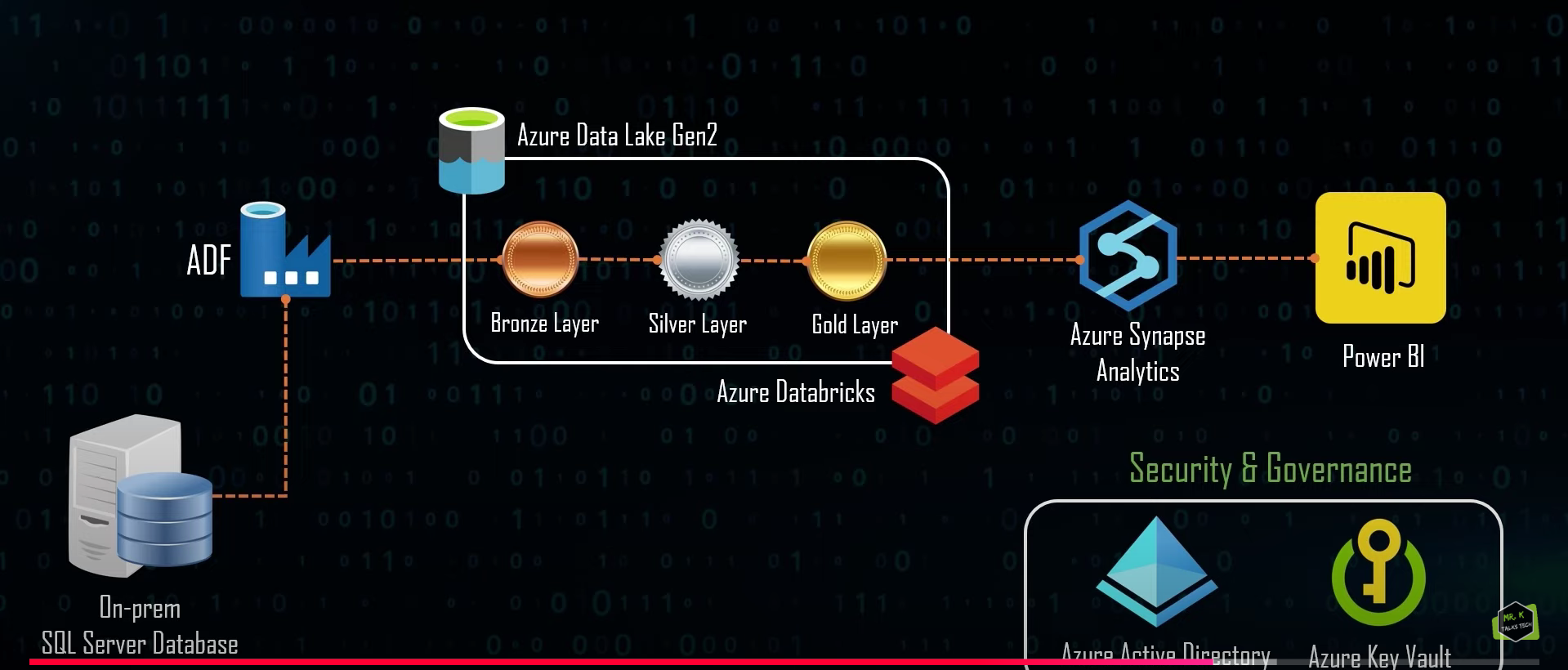
5. Azure Key vault

6. Azure Active Directory (AAD)

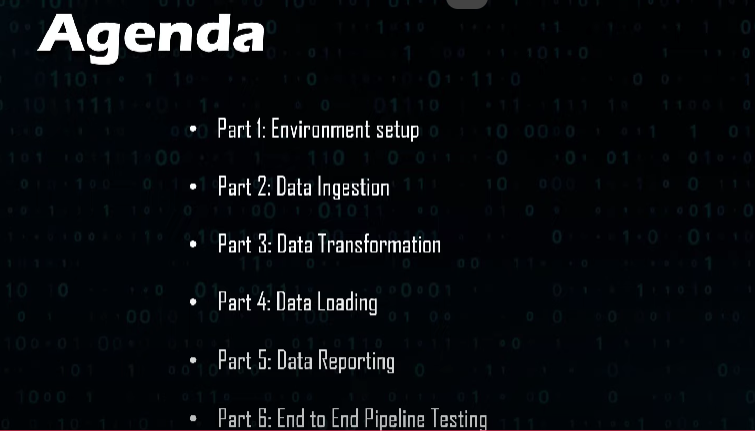
7. Microsoft Power BI



**Project Architecture**



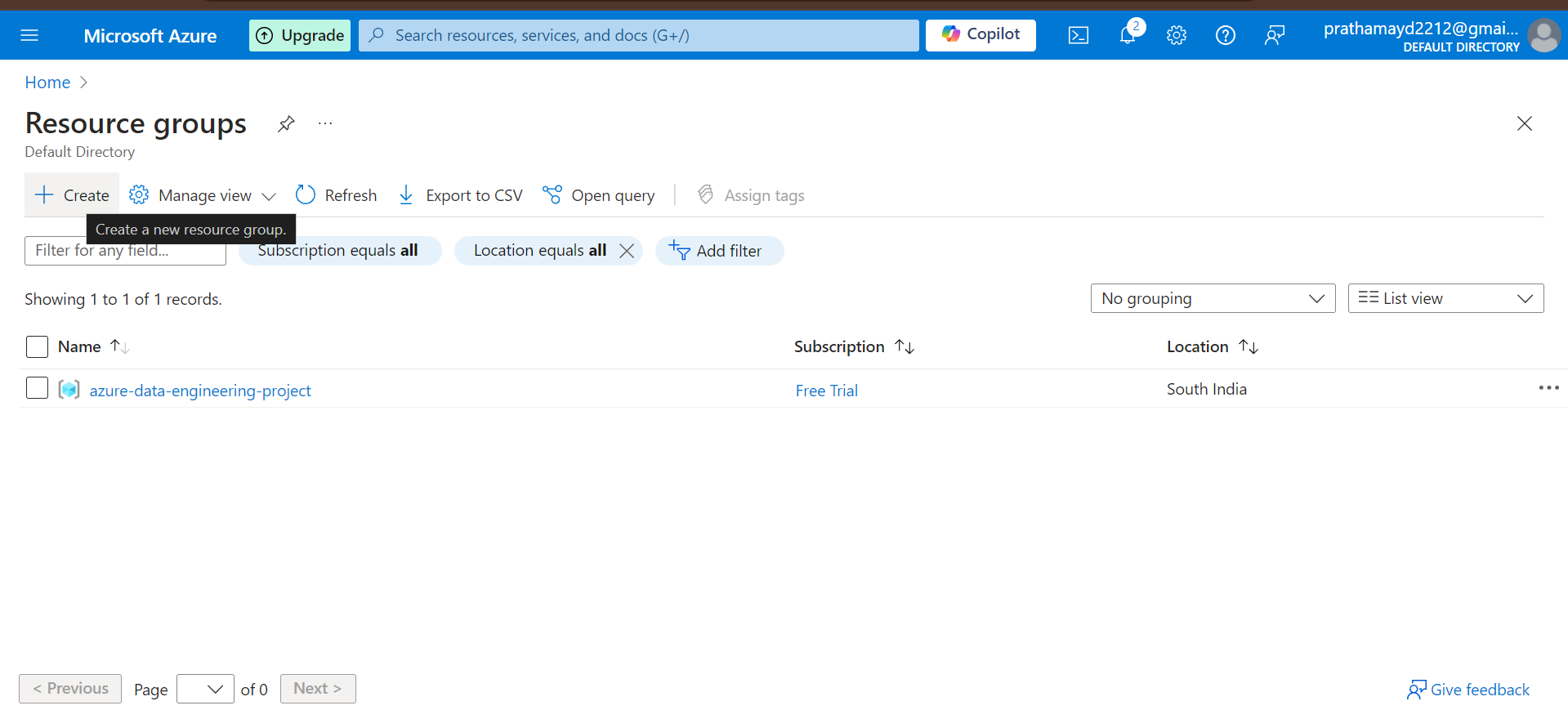
1. Data ingestion
2. Transformation of data
3. Creating database and loading the data in the database
4. Using power Bi for creating reports
5. Everything will be automated.



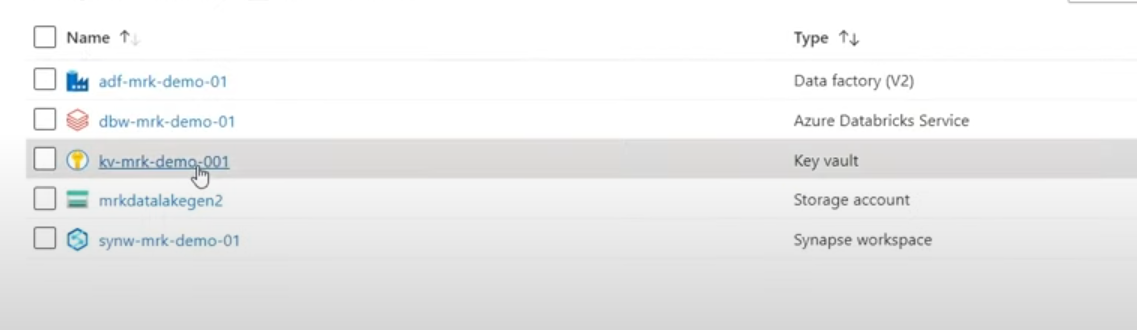
**1-Environment Setup**

Firstly we need to create Resource Group under which all of the resources can be created to execute the project.

Search for resource groups and create a resource group for the project.

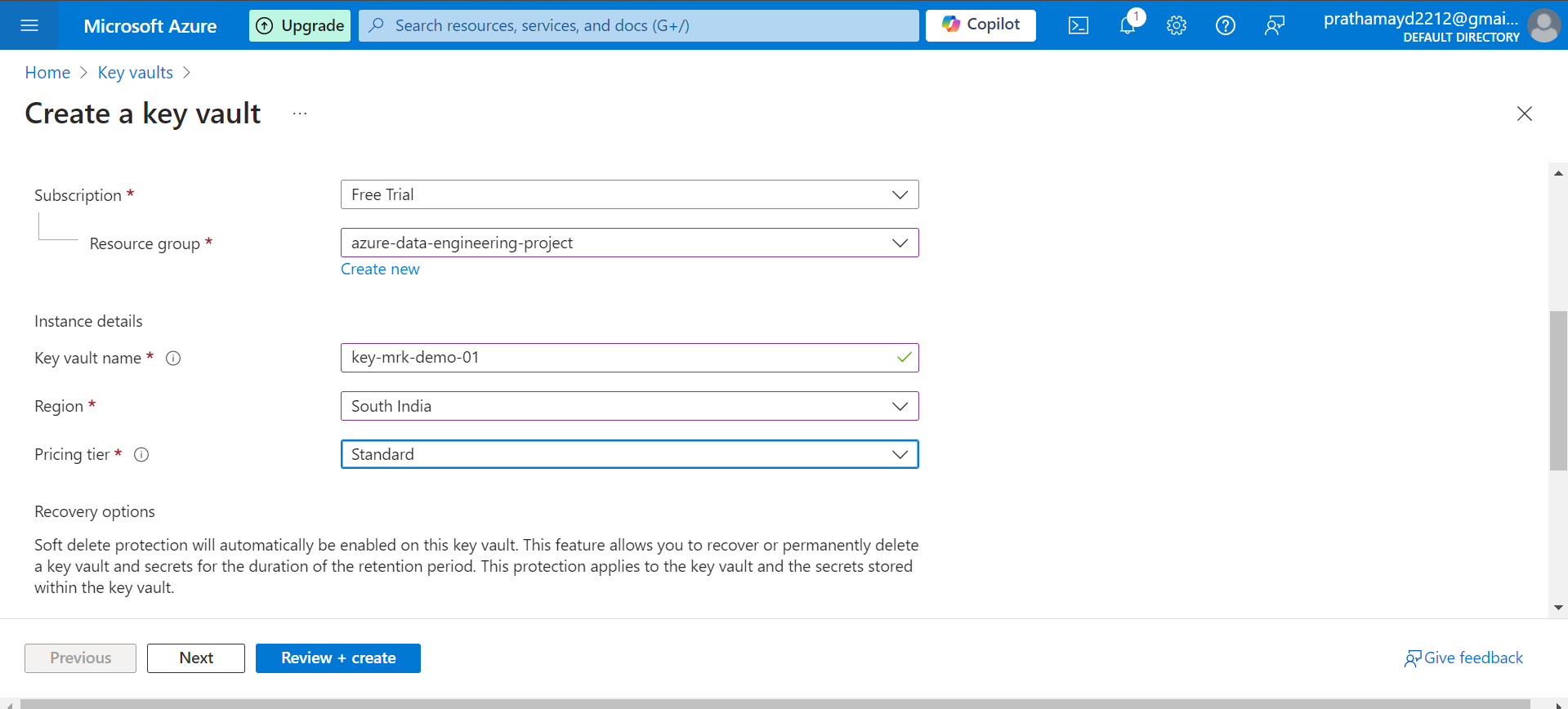


We have to create these resources under the resource group that was created earlier:



We will start with creating a key vault resource.

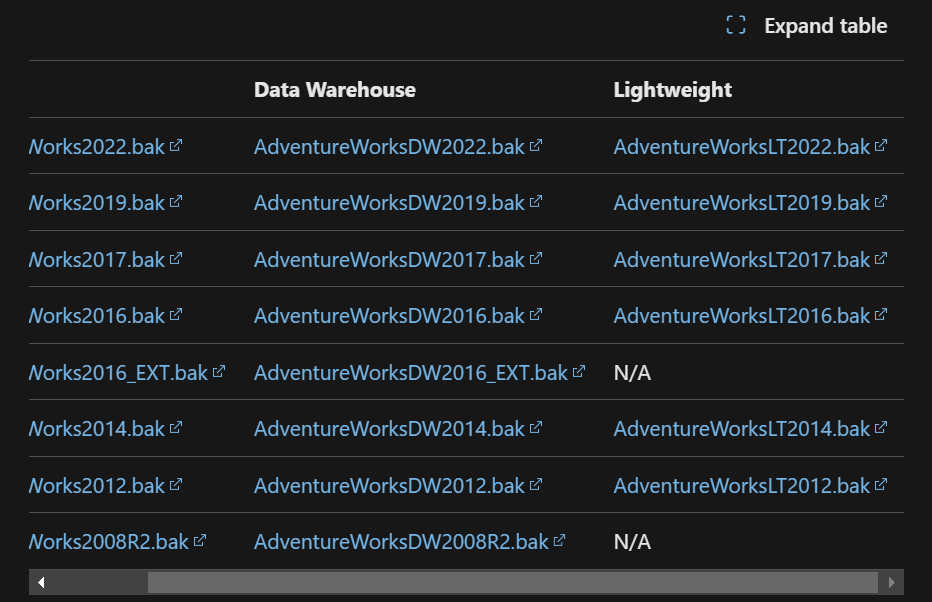
Search for key vaults, and create the key vault under the resource group that we created initially.



Once created then we need to create store our SQL Server username and password in this key vault’s Secrets section so that the username and password are stored in an encrypted format and will remain safe from everyone.

Now we will download the dataset from the link provided below:  
<https://learn.microsoft.com/en-us/sql/samples/adventureworks-install-configure?view=sql-server-ver16&tabs=ssms>

I have downloaded 2017 data.



Now restore the data into your database so that it can be used from ssms.

Right click on the database🡪Restore Database🡪From device🡪Browse the file🡪 select and click on restore a new database with the dataset will be created.

Store the downloaded file the path given below:  
C:\Program Files\Microsoft SQL Server\MSSQL16.SQLEXPRESS\MSSQL\Backup

FILE:



🡪once the database is restored

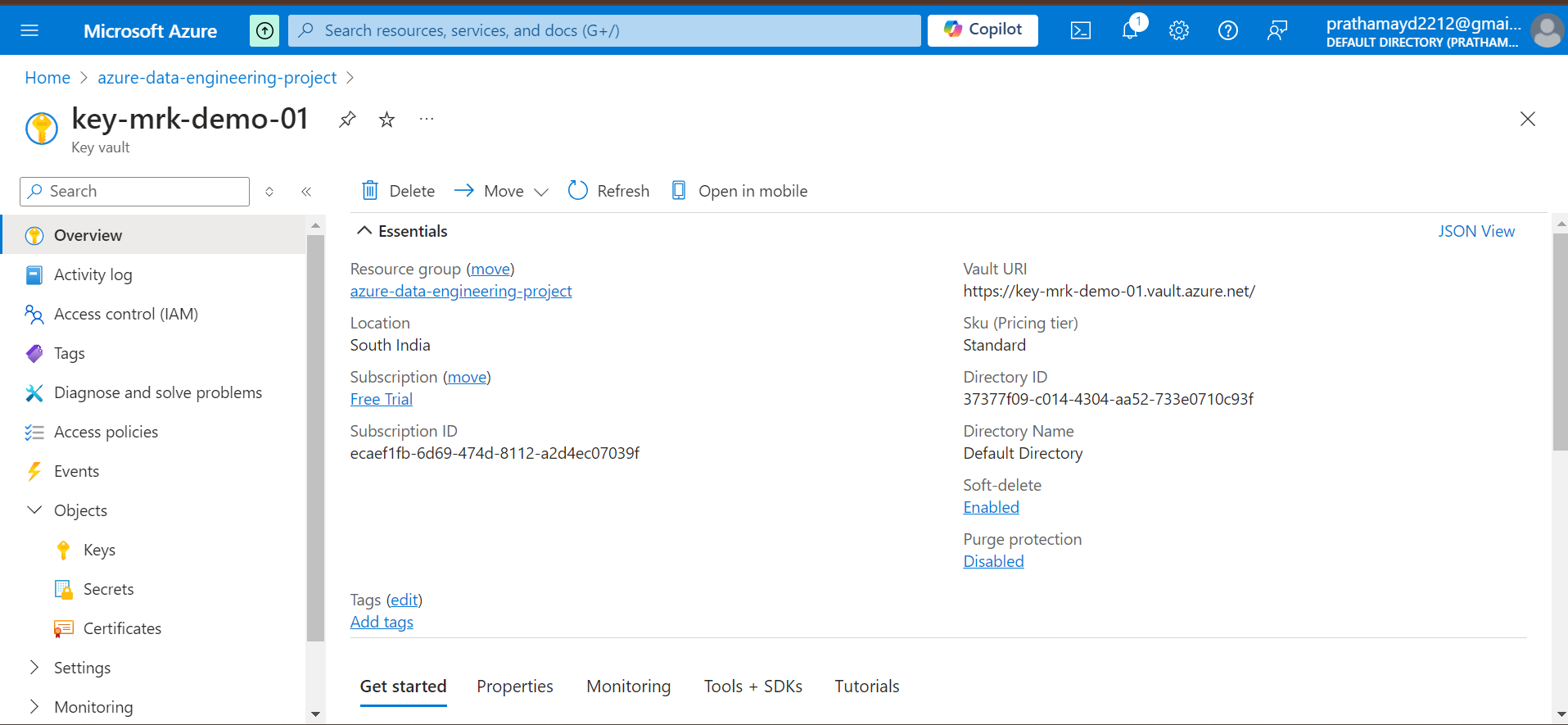
You need to look for you login and password and store that as the secret to create the access.

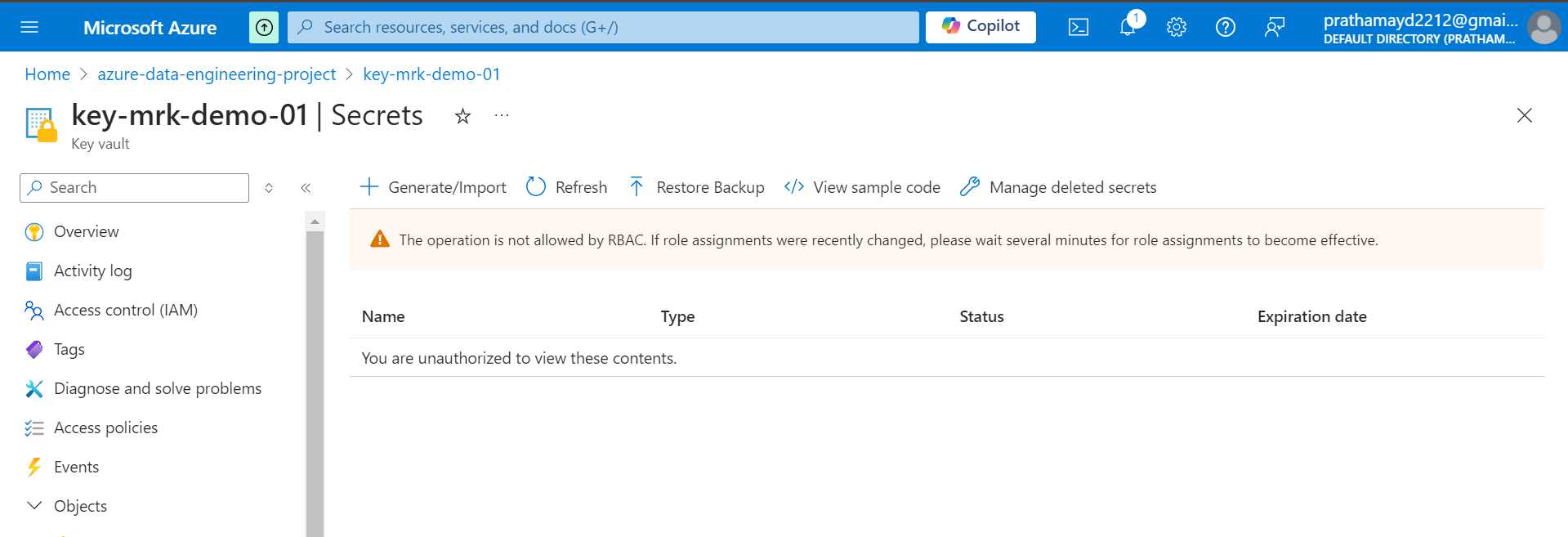
**User: azurepratham**

**Password: pvk18000**

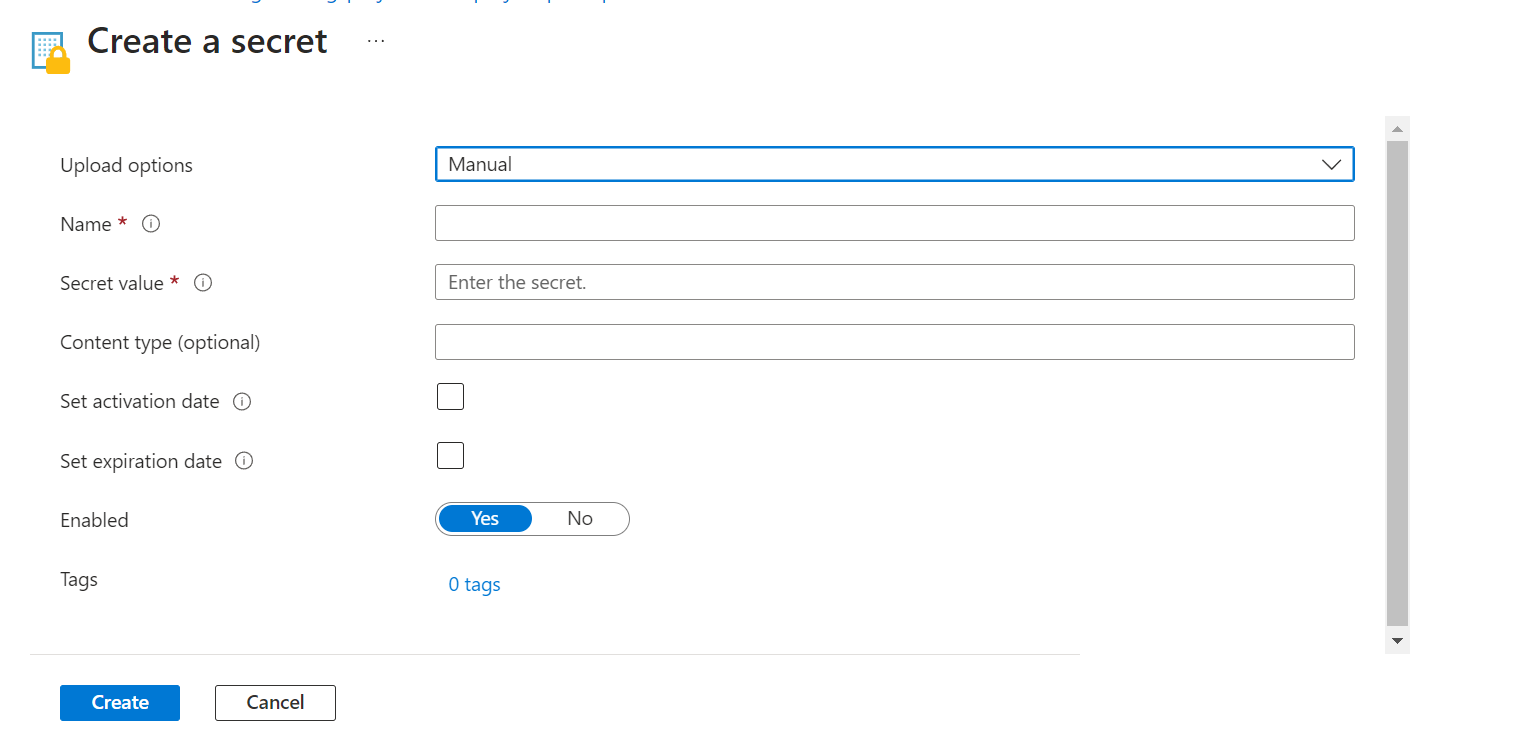
🡪Give encryption to the user and password

Navigate to key vault resource in your resource group and go to secrets.



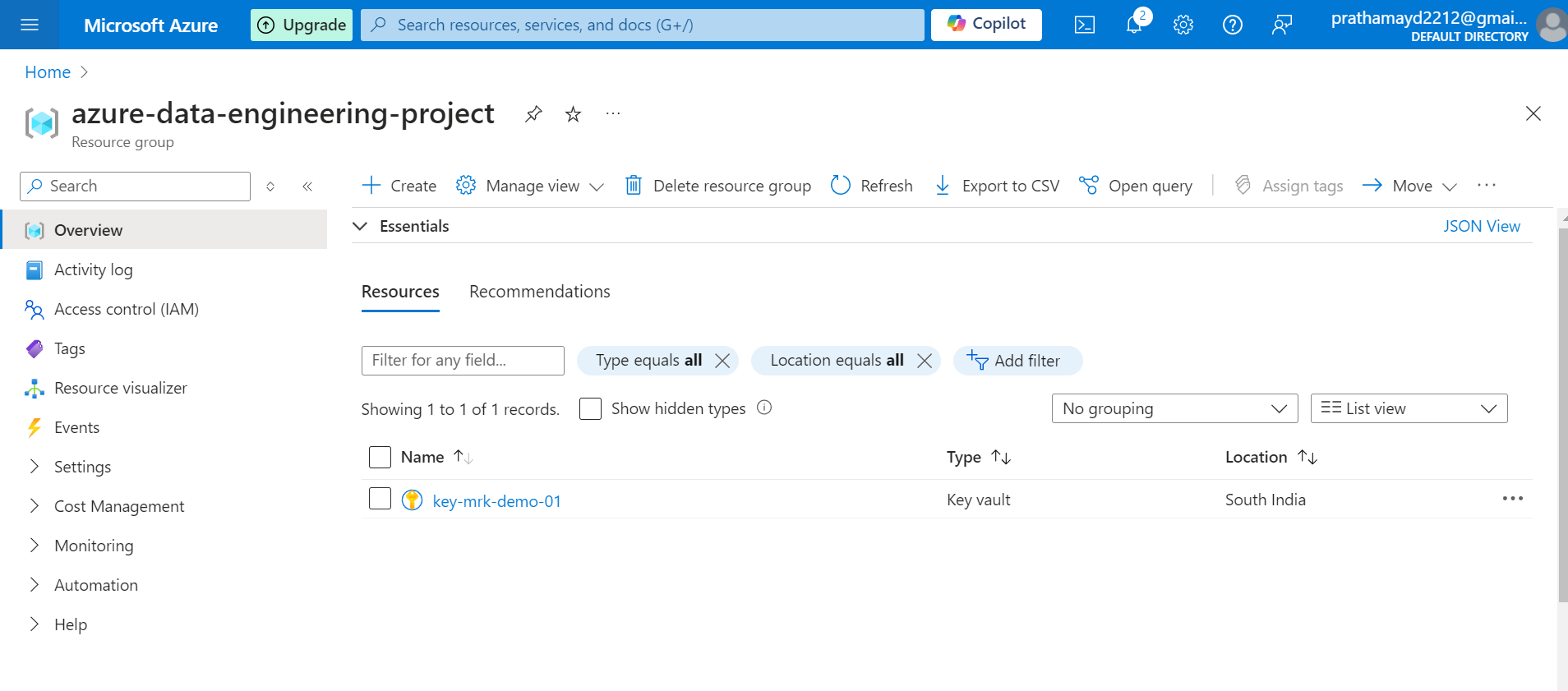


If you do not have access to create secrets, you can navigate to IAM roles🡪add role assignment🡪provide Key Vault Administrator role and create.

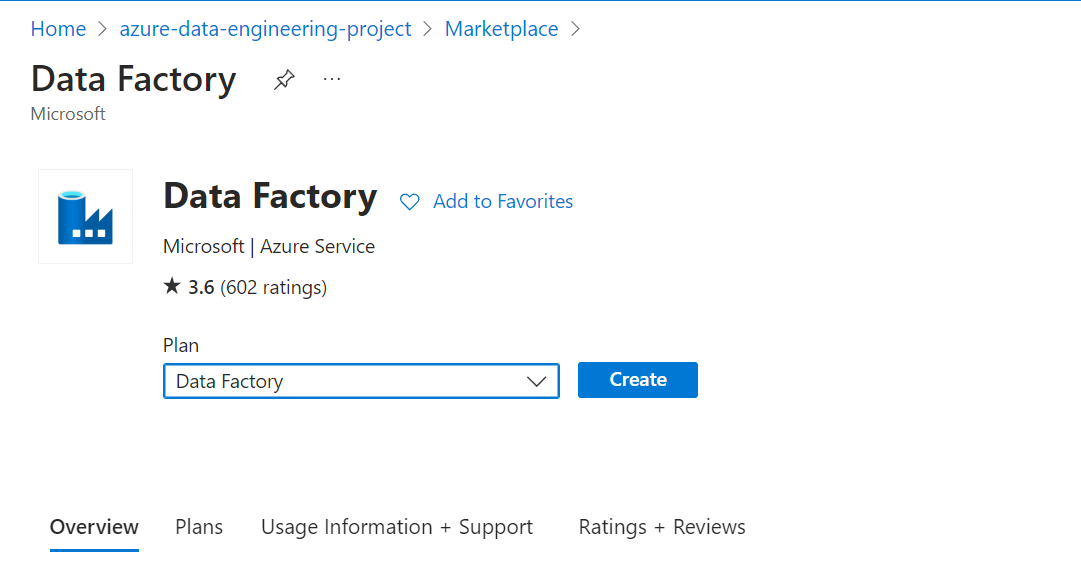


Create two secrets, one for username and other one for password.

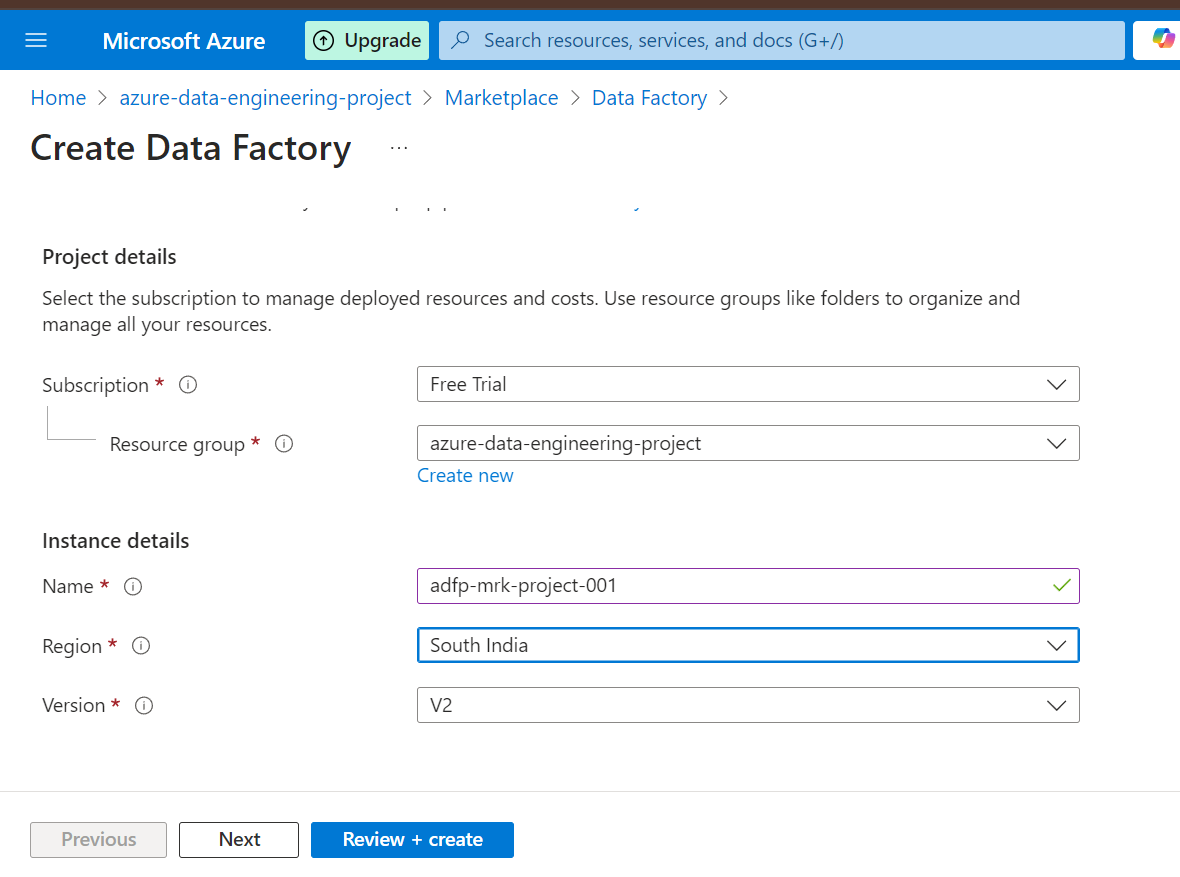
🡪Creating Data factory resource under our resource group.



And search for Data Factory and select the data factory. And click on Create.



🡪

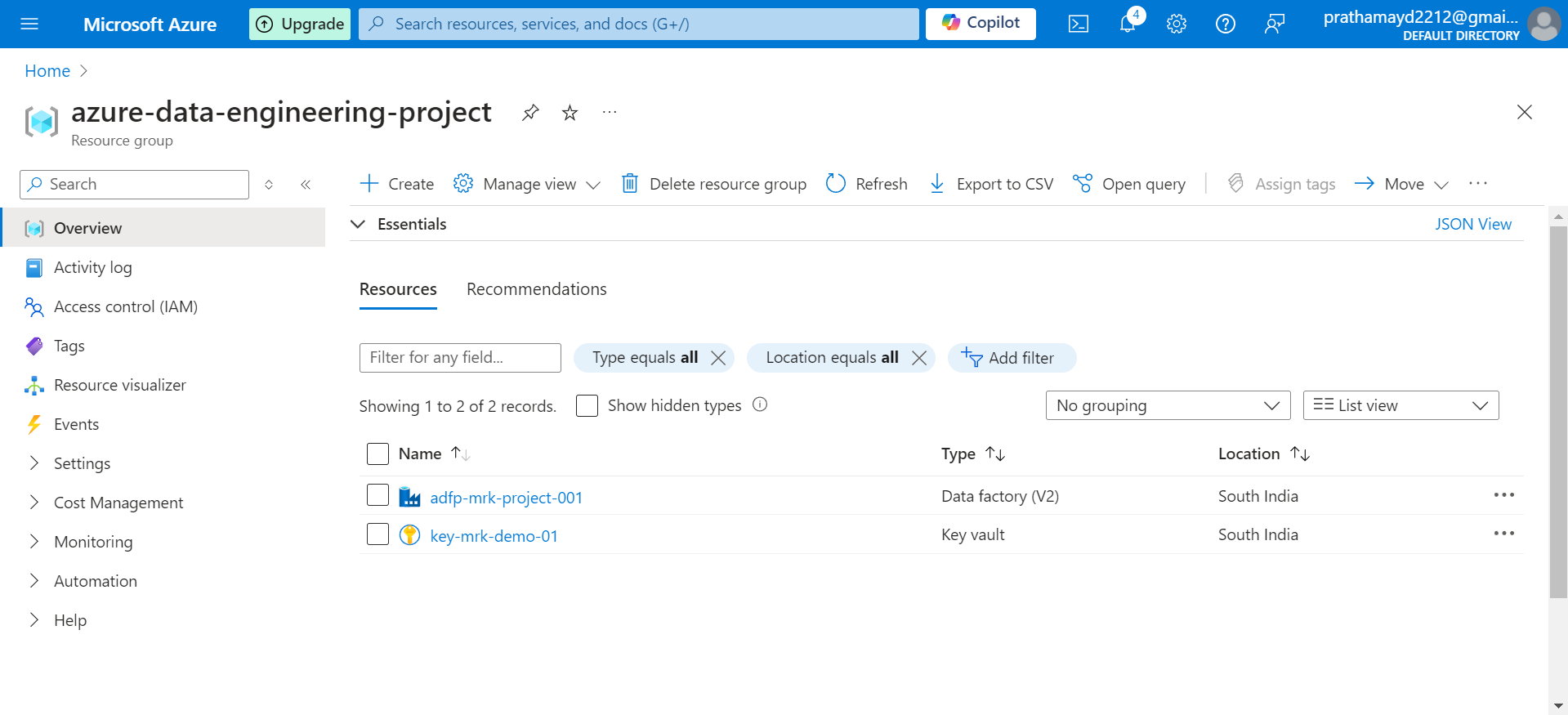


🡪Network

Set it as Private endpoint

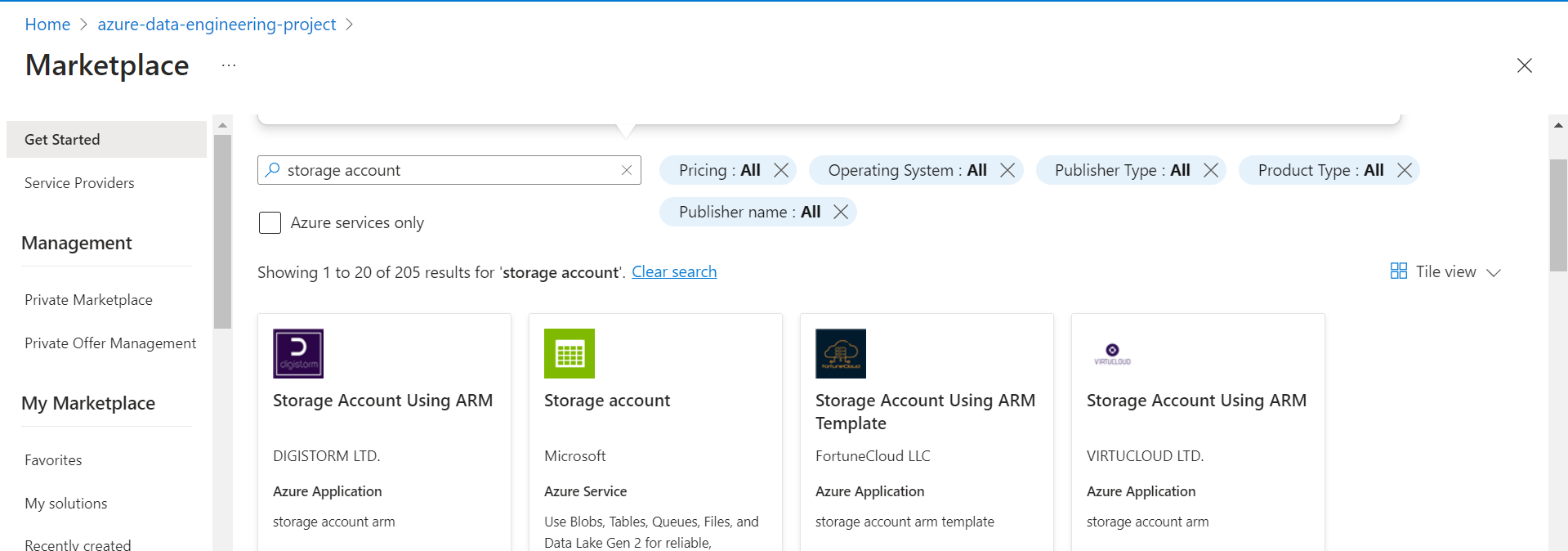
Keep the other options unchanged and click on review + create

Once done, the resource will be reflected in your resource group.

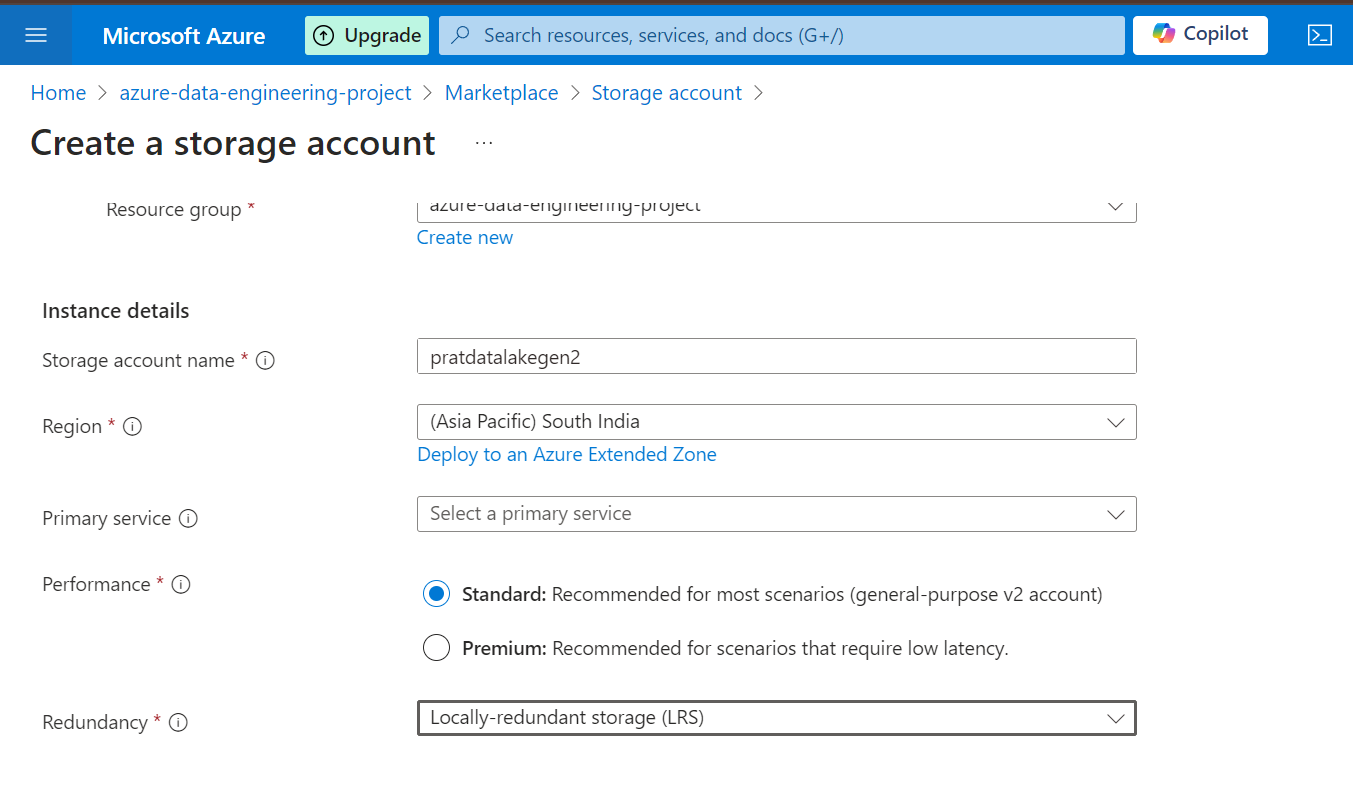


🡪Creating Storage Account resource in our resource group.

The process is same, we need to open the resource group🡪create🡪search storage account and create the account

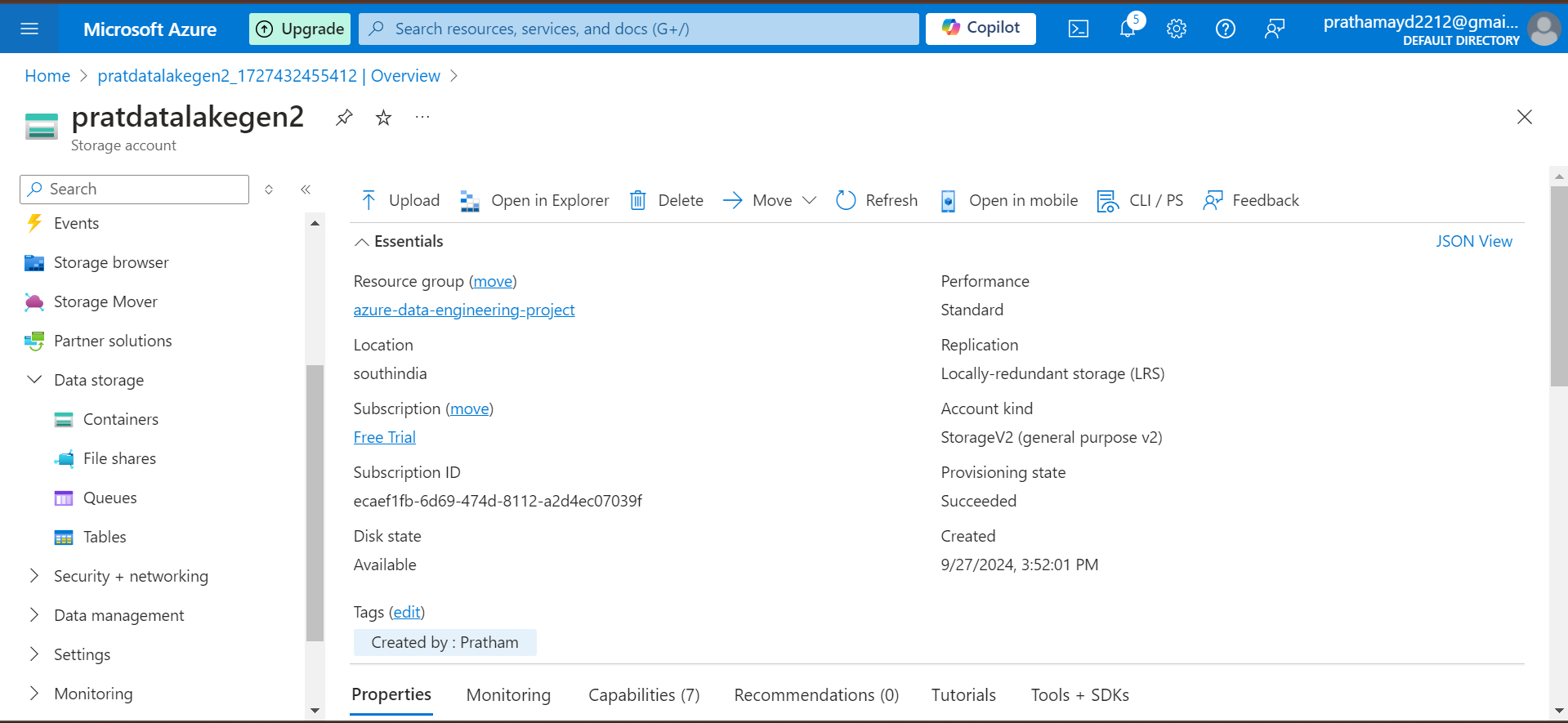


🡪Basics



Keep everything else as default.  
you can choose the primary service as : **Azure Blob Storage or Azure Data Lake Storage Gen 2**

And review + Create the resource.  
Once created, we need to open the resource and create a container in it.



Click on add container, give a container name and create a container.

(NOTE: Container name should only contain small letter words and nothing else is accepted.)

NOTE(

CREATE THE AZURE SYNAPSE WORKSPACE AT THE BEGINNING OF THE PROJECT, WHILE CREATING THE SYNAPSE RESOURCE YOU WILL GET AN OPTION TO CREATE A STORAGE ACCOUNT AND A CONTAINOR, USE THAT TO STORE YOUR BRONZE, SILVER AND GOLD CONTAINERS AND ONLY THEN THE CONTAINER WILL LINK TO THE AZURE SYNAPSE WORKSPACE OTHERWISE YOU WILL HAVE TO CREATE A LINKED SERVICE SEPARATELY.