

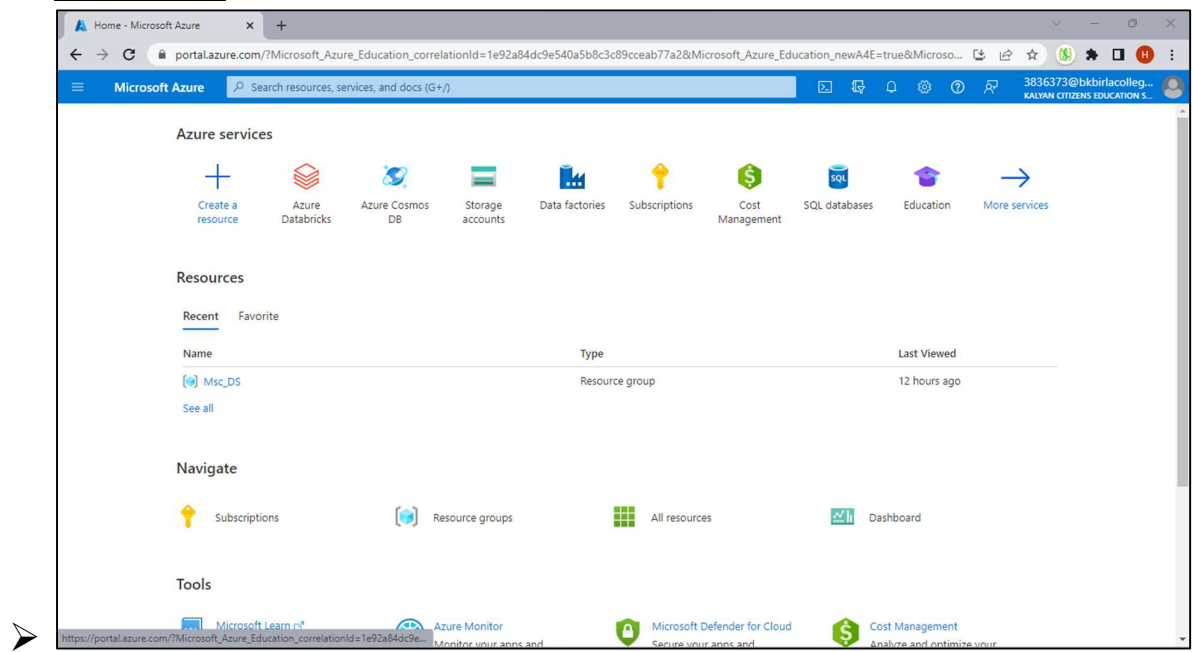
Data on cloud

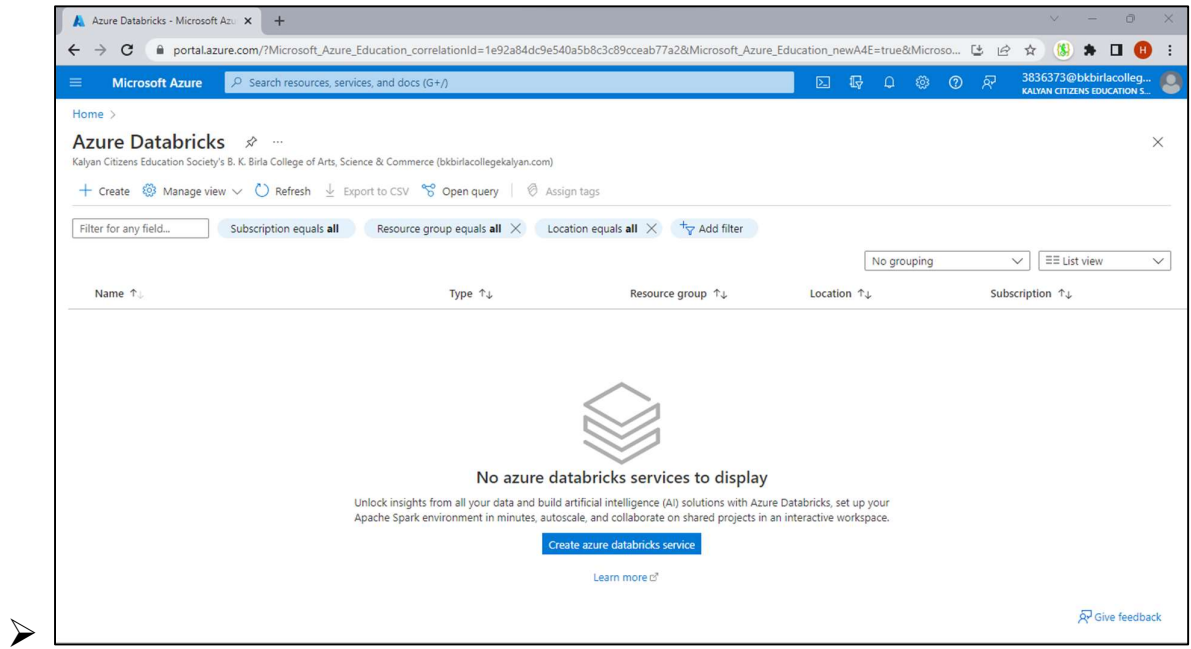
[Practical 4 & 5]

2 Dec 2022

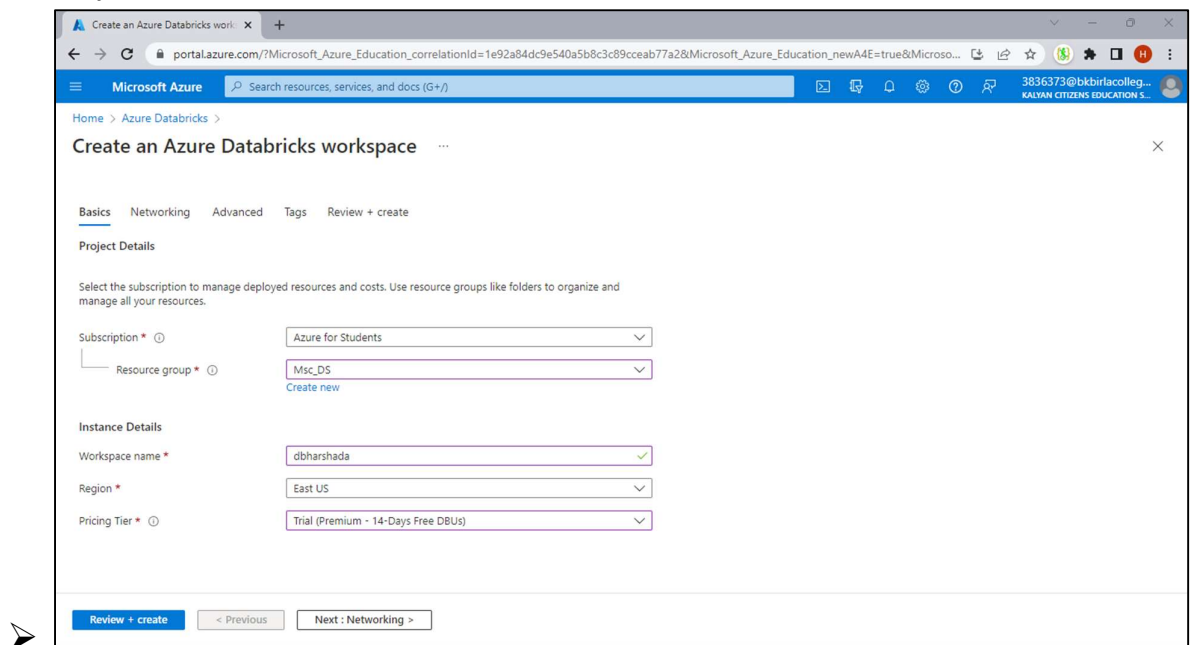
P.no.[4] Create Azure databricks

- Go to the home page of Microsoft azure > click on Azure Databricks

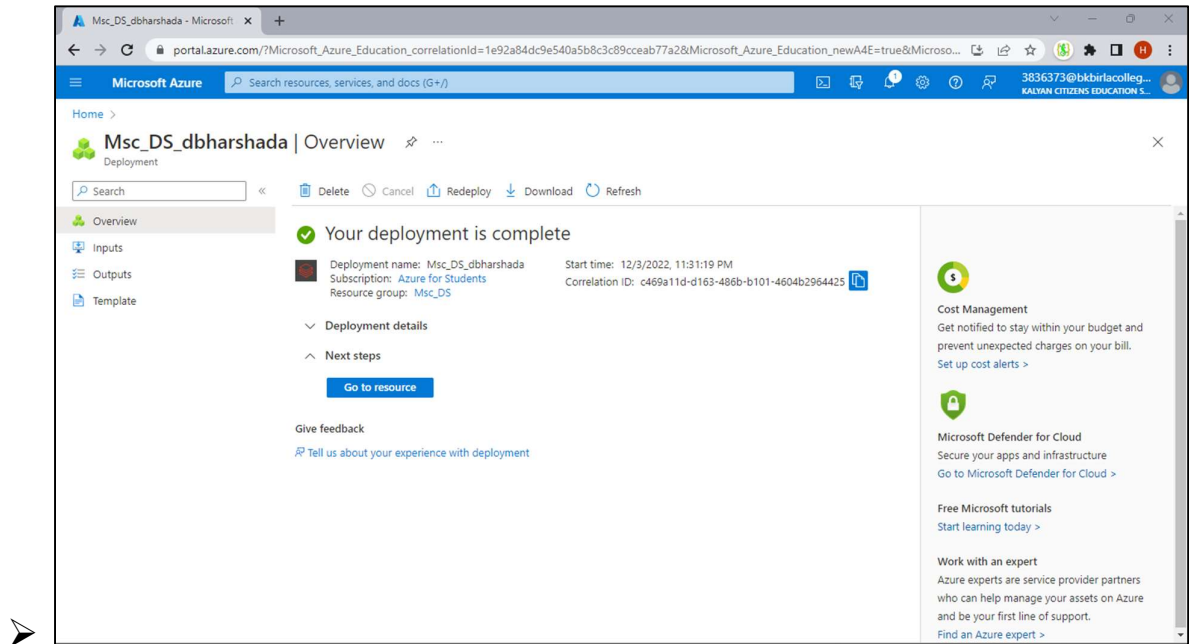




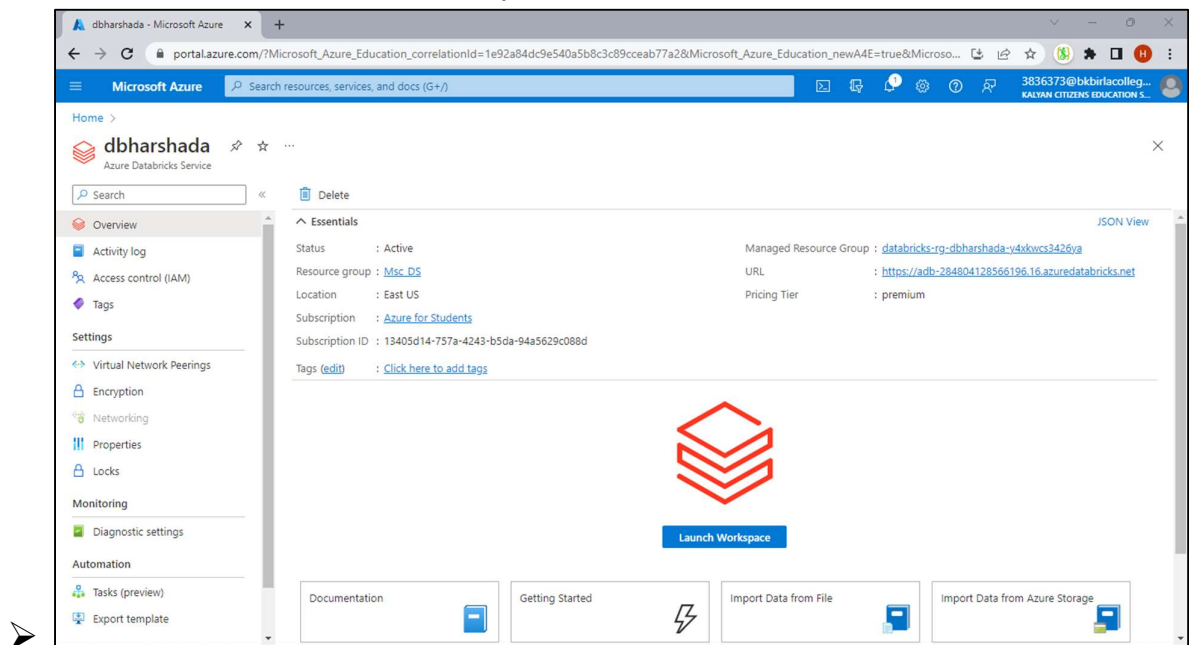
- Click on create > select the resource group (Msc_DS) > In instance details – Workspace name : dbharshada (dbfirstname) > click on Review+create > pricing tier : select -Trial for 14 days.



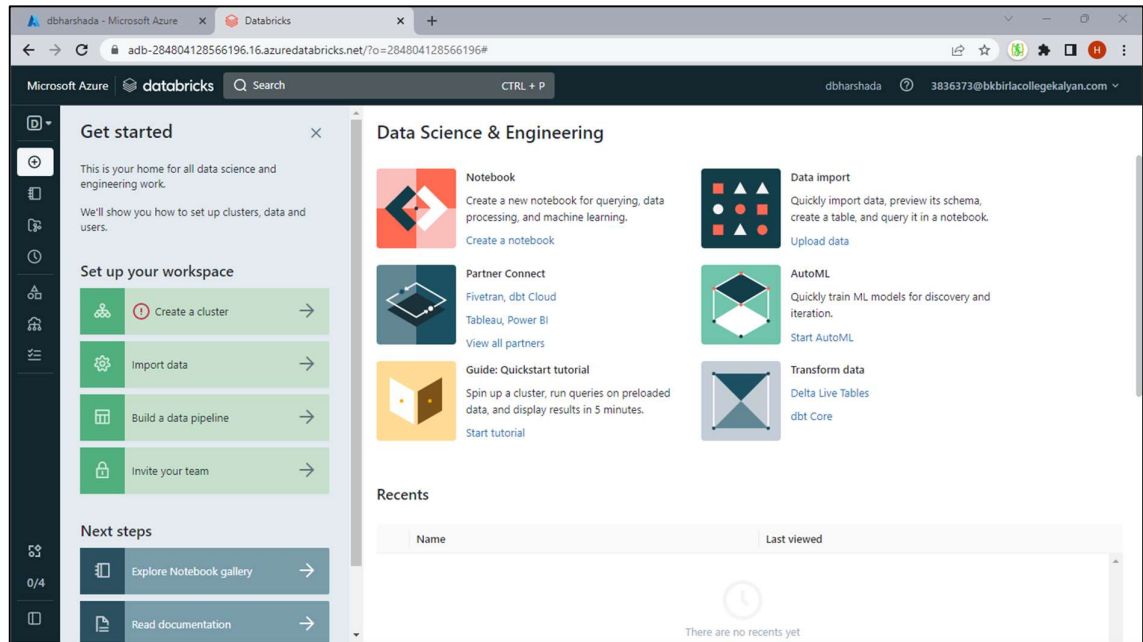
- Then we get..



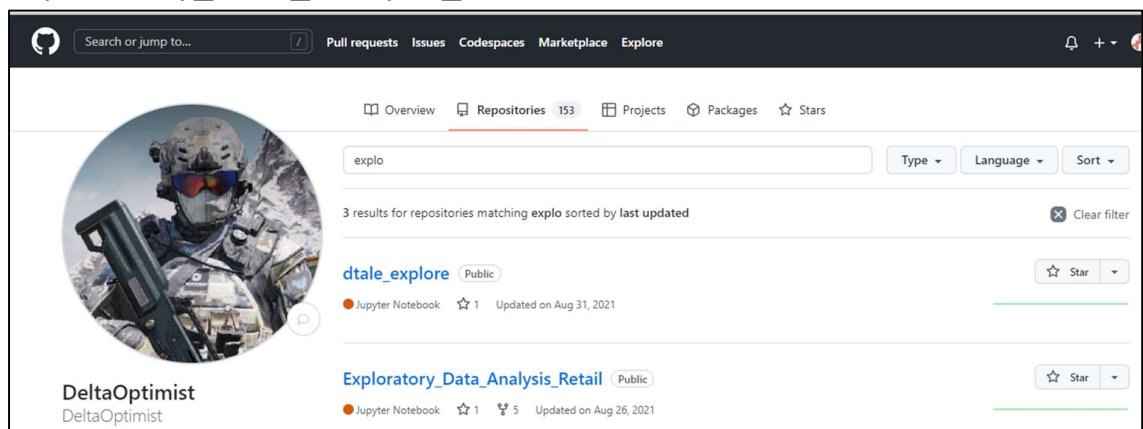
- Come back to Home page and click on dbharshada(azure databricks) > Launch Workspace .



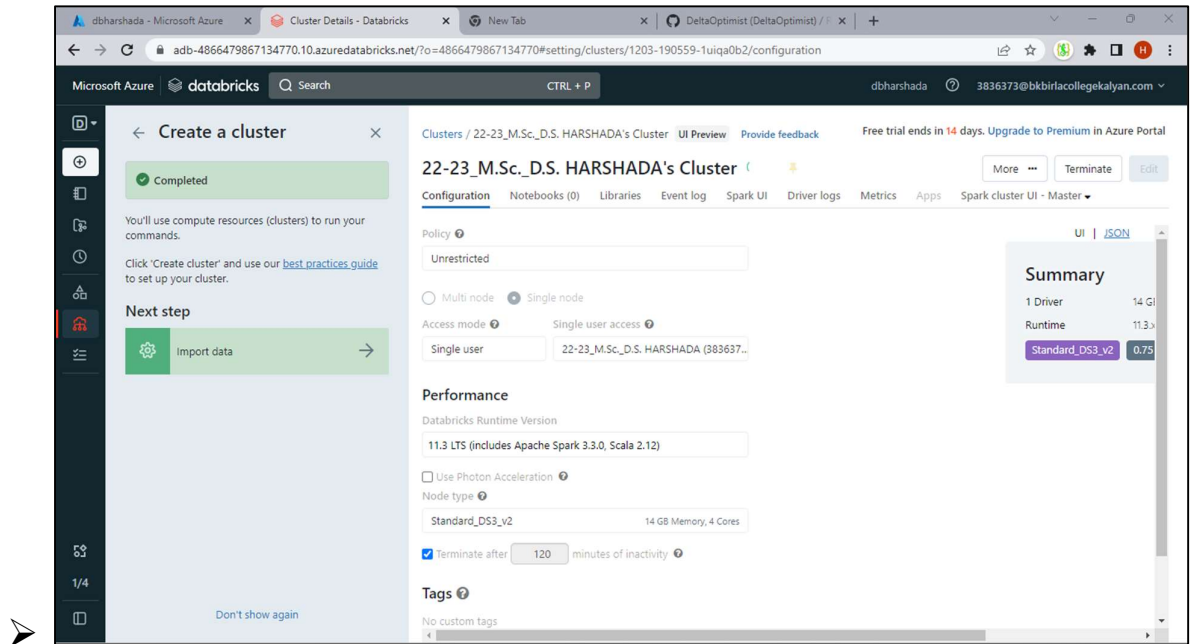
- When we launch workspace then we get some questions or ask some about project, it is not important just skip or close .
- Then get...



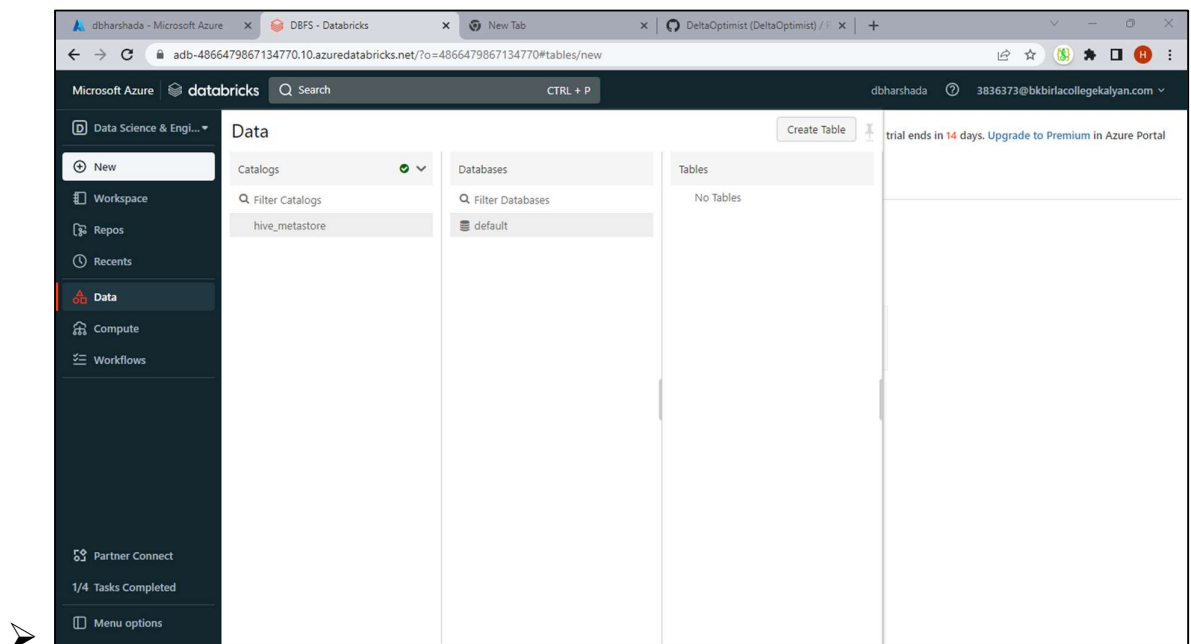
- Now go on new tab to download and extract the data set from deltaoptimist github, which is Exploratory_Data_Analysis_Retail.

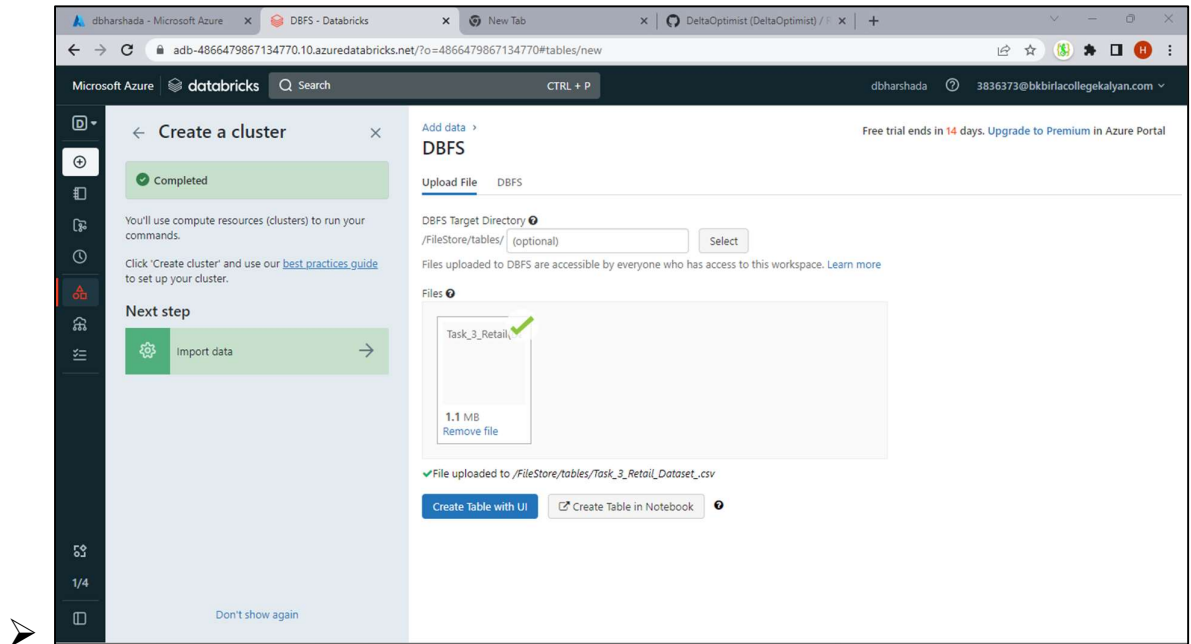


- Next create cluster --> compute > create compute>all details keep as it is > click below create cluster.
If there has come any error to create cluster the switch the Multi Node to – single Node.



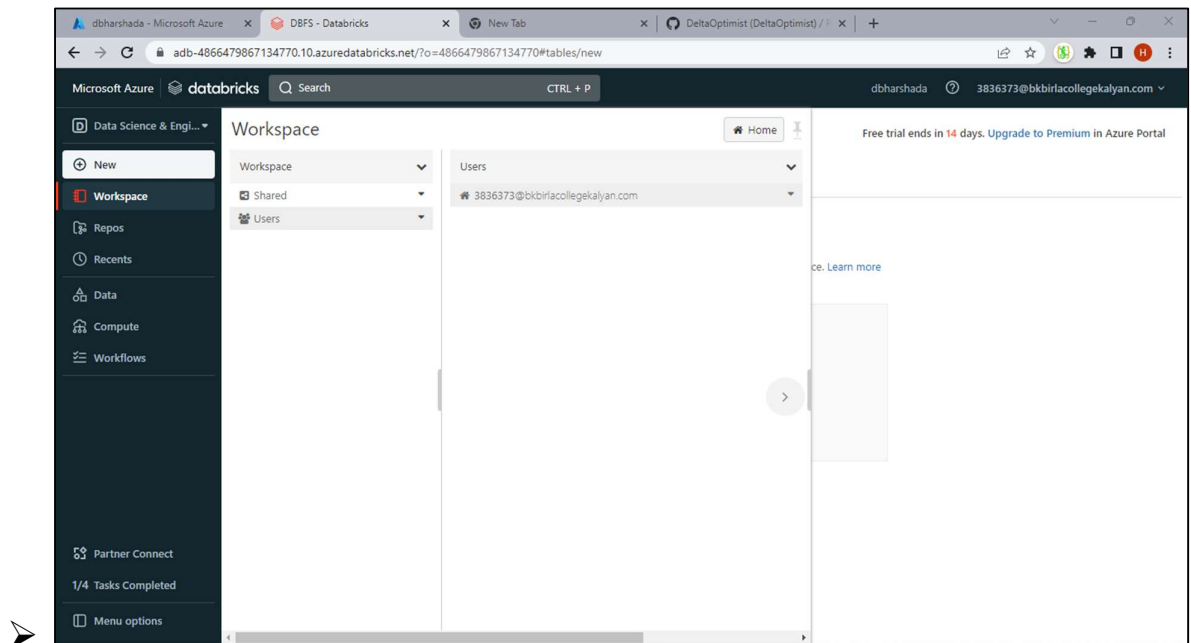
- Click on Data > then click on create table > upload the file which we download and extract from the GitHub >

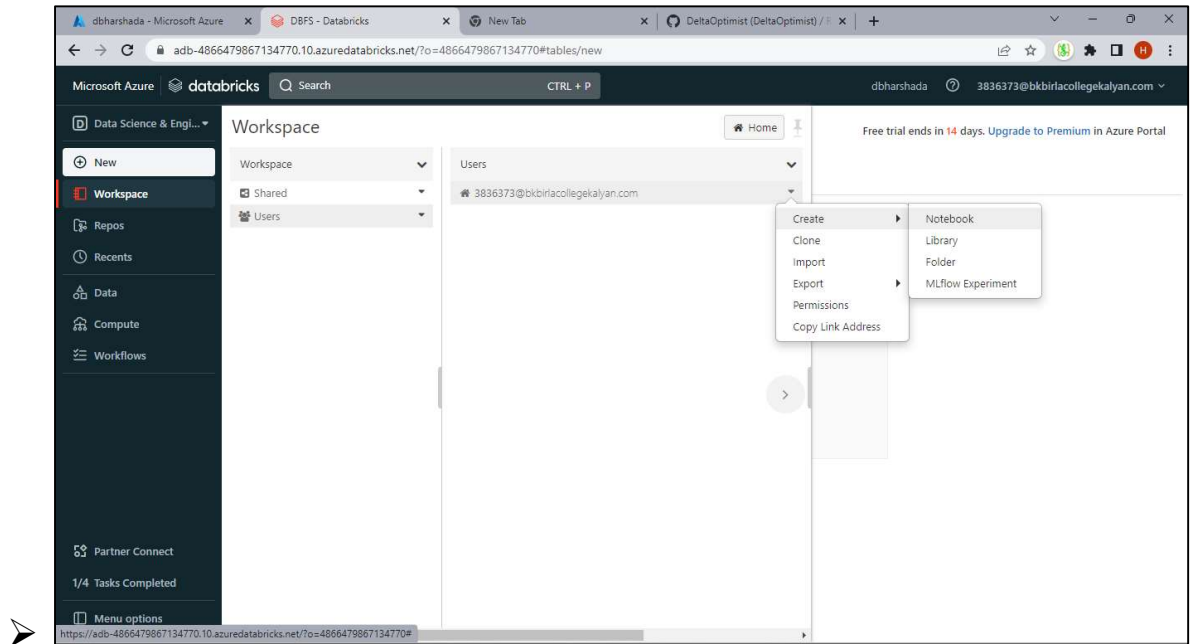




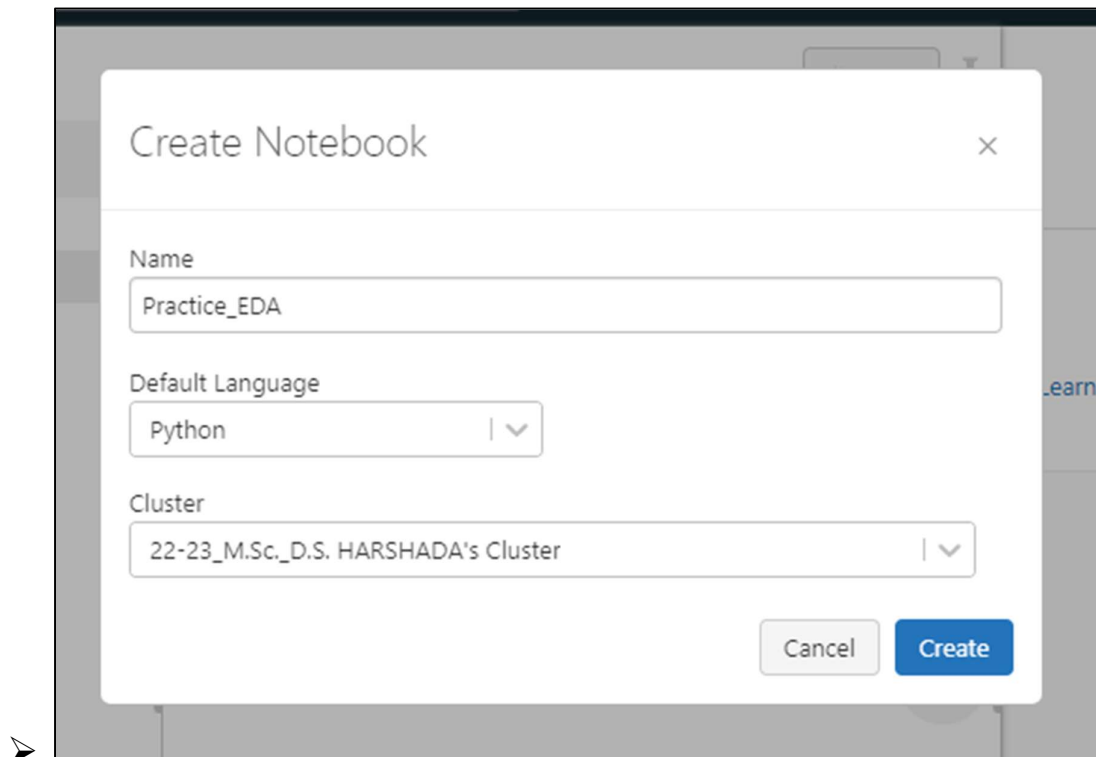
➤ Then copy the path where we upload the file. `/FileStore/tables/Task_3_Retail_Dataset.csv`

➤ Go on workspace > select User > click on your user id > create > notebook.

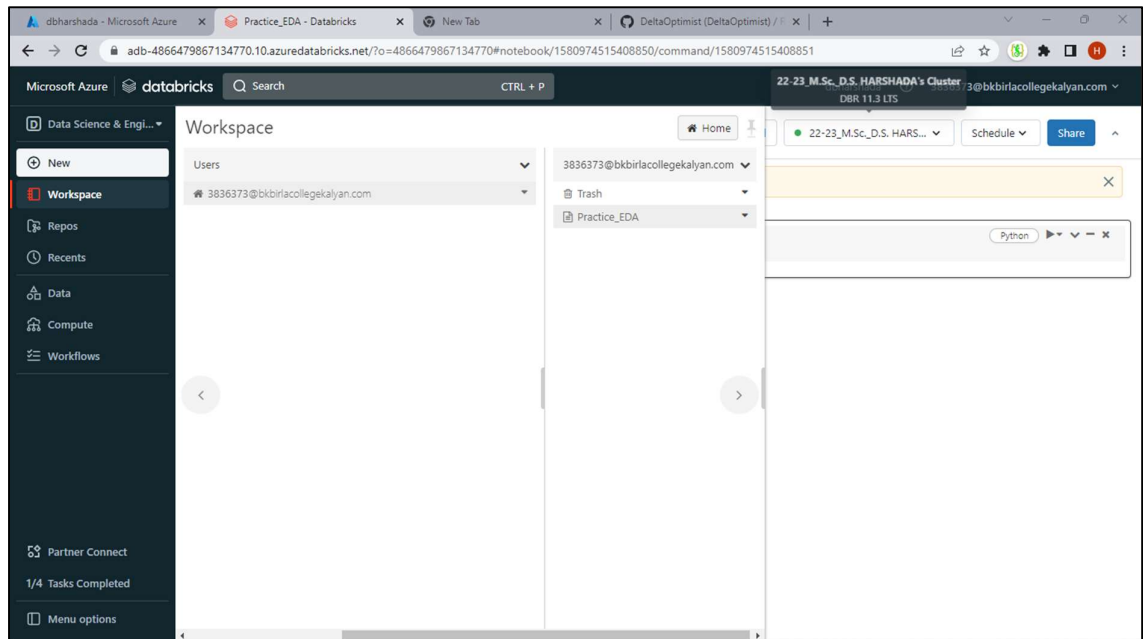




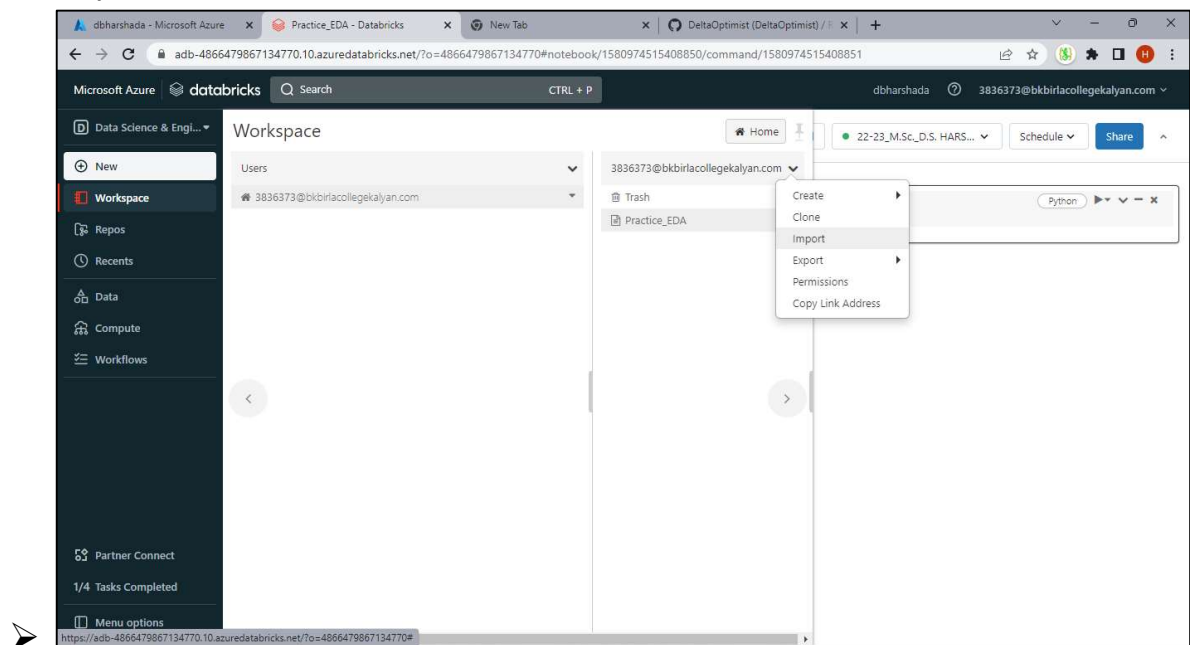
- Create Notebook > Give the name of notebook > python > select cluster > click on create.



- Again select workspace > click on Practice_EDA – to create the python notebook.



- import the data set on notebook.



- The path : */dbfs/FileStore/tables/Task_3_Retail_Dataset.csv*

➤ then get ...

The screenshot shows a Databricks notebook interface. The notebook is titled "Exploratory_Data_Analysis_Retail" and is running on a cluster named "22-23_M.Sc._D.S. HARSHADA's Cluster". The code in the notebook is as follows:

```
1 # Loading dataset
2 df_data = pd.read_csv('/dbfs/FileStore/tables/Task_3_Retail_Dataset_.csv')
3 # Printing top 5 rows
4 df_data.head()
```

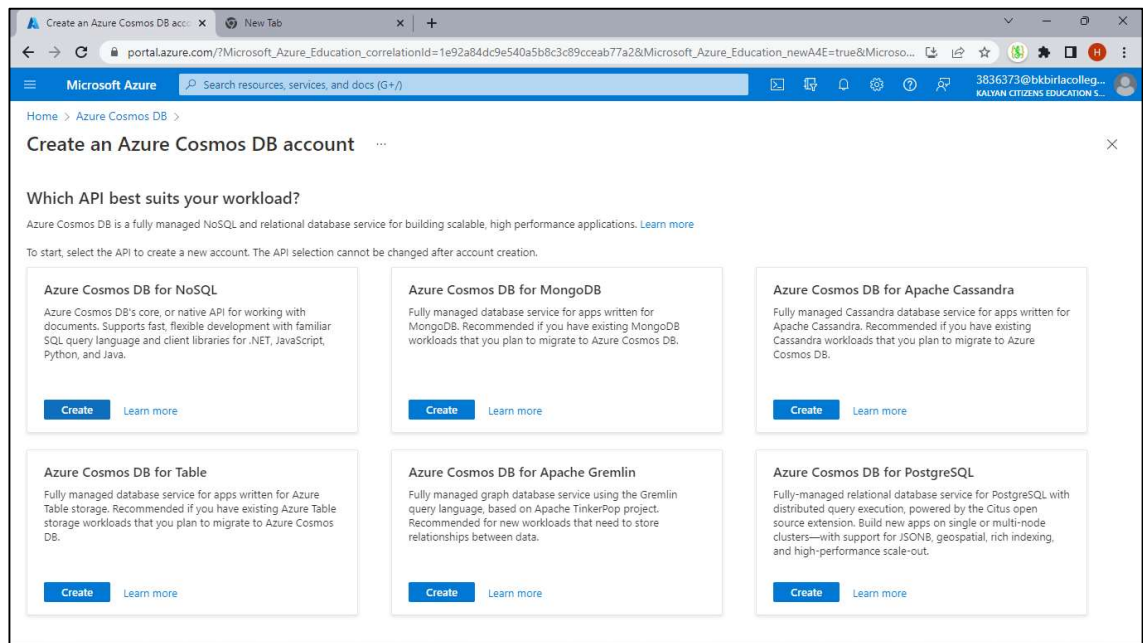
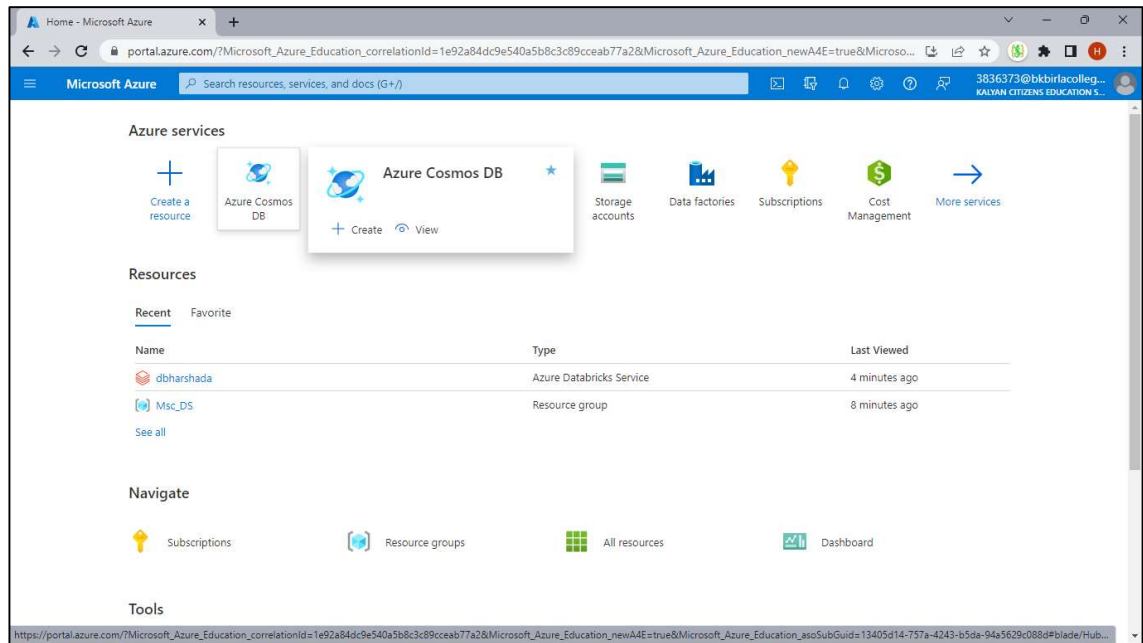
The output of the code is a pandas DataFrame with the following columns: Ship Mode, Segment, Country, City, State, Postal Code, Region, Category, Sub-Category, Sales, Quantity, Discount, and Profit. The DataFrame contains 5 rows of data:

	Ship Mode	Segment	Country	City	State	Postal Code	Region	Category	Sub-Category	Sales	Quantity	Discount	Profit
0	Second Class	Consumer	United States	Henderson	Kentucky	42420	South	Furniture	Bookcases	261.9600	2	0.00	41.9136
1	Second Class	Consumer	United States	Henderson	Kentucky	42420	South	Furniture	Chairs	731.9400	3	0.00	219.5820
2	Second Class	Corporate	United States	Los Angeles	California	90036	West	Office Supplies	Labels	14.6200	2	0.00	6.8714
3	Standard Class	Consumer	United States	Fort Lauderdale	Florida	33311	South	Furniture	Tables	957.5775	5	0.45	-383.0310
4	Standard Class	Consumer	United States	Fort Lauderdale	Florida	33311	South	Office Supplies	Storage	22.3680	2	0.20	2.5164

P.no.[5]Create cosmosDB container, insert data records in containers and delete the records using TTL(Time to Live) approach.

➤ Home page > click on Azure cosmosDB > create new Azure cosmosDB account > select Azure cosmos DB for NoSQL > select resource group > Type the Account name : cosmosdb14 > click

on create.



Create Azure Cosmos DB Account - Azure Cosmos DB for NoSQL

Basics Global Distribution Networking Backup Policy Encryption Tags Review + create

Azure Cosmos DB is a fully managed NoSQL and relational database service for building scalable, high performance applications. Try it for free, for 30 days with unlimited renewals. Go to production starting at \$24/month per database, multiple containers included. [Learn more](#)

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 for Students
Resource Group * Msc_DS
[Create new](#)

Instance Details
Account Name * cosmosdb14
Location * (US) West US
Capacity mode ☒ Provisioned throughput ☐ Serverless
[Learn more about capacity mode](#)

With Azure Cosmos DB free tier, you will get the first 1000 RU/s and 25 GB of storage for free in an account. You can enable free tier on up to one account per subscription. Estimated \$64/month discount per account.

Apply Free Tier Discount ☒ Apply ☐ Do Not Apply

[Review + create](#) [Previous](#) [Next: Global Distribution](#)

➤ Deployment is complete.

Microsoft.Azure.CosmosDB-20221204194808 | Overview

Deployment

Search [Delete](#) [Cancel](#) [Redeploy](#) [Download](#) [Refresh](#)

✓ Your deployment is complete

Deployment name: Microsoft.Azure.CosmosDB-20221204... Start time: 12/4/2022, 7:48:14 PM
Subscription: Azure for Students Correlation ID: 72c5ac11-f16f-4abb-acc6-ca55ac3e92c1
Resource group: Msc_DS

Deployment details
Next steps
[Go to resource](#)

Give feedback
[Tell us about your experience with deployment](#)

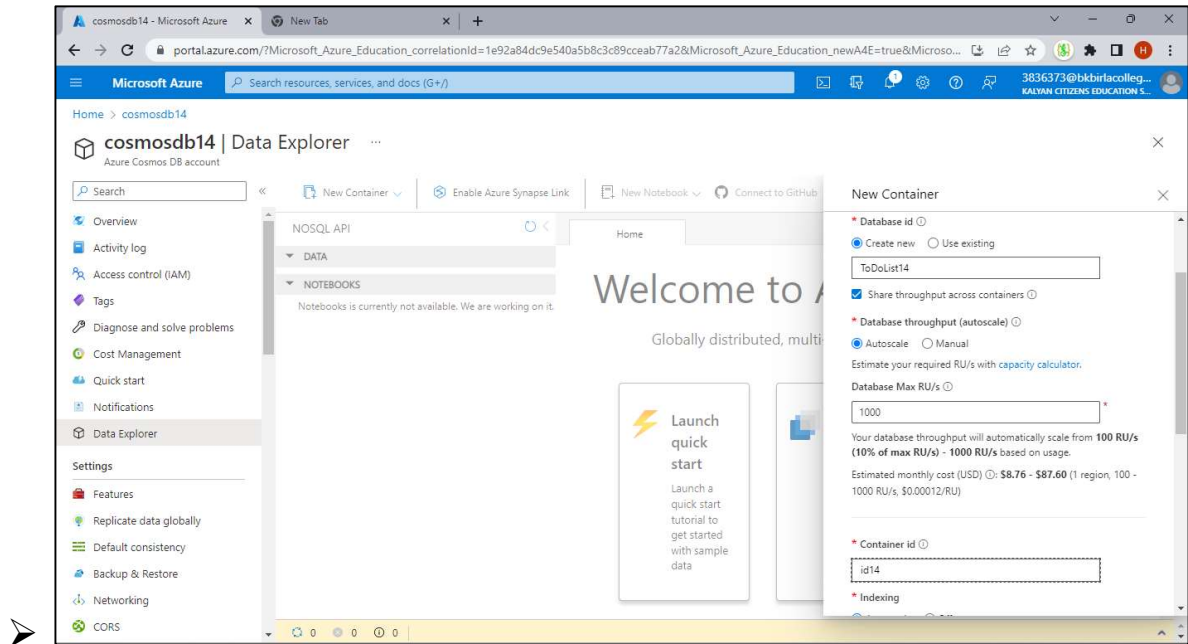
Cost Management
Get notified to stay within your budget and prevent unexpected charges on your bill.
[Set up cost alerts >](#)

Microsoft Defender for Cloud
Secure your apps and infrastructure
[Go to Microsoft Defender for Cloud >](#)

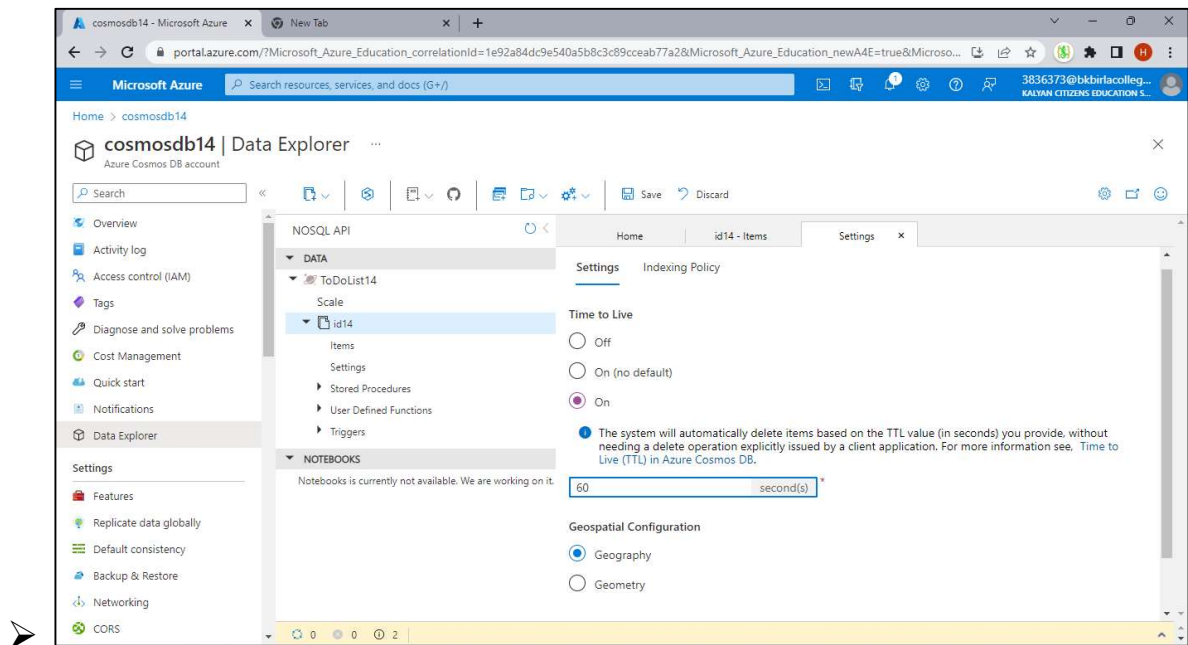
Free Microsoft tutorials
[Start learning today >](#)

Work with an expert
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.
[Find an Azure expert >](#)

➤ Home > click on cosmosDB > add container > fill the details > click on ok .



- Click on ToDoList14 > id14 > items > go on items setting edit time to live with On > set time in second > Save .



- Comes on id14-items > click on New item > type query > save > add new Query > Execute Query > it automatically delete after 60 sec.

cosmosdb14 - Microsoft Azure x New Tab x +

portal.azure.com/7Microsoft_Azure_Education_correlationId=1e92a84dc9e540a5b8c3c89cceab77a2&Microsoft_Azure_Education_newA4E=true&Microso...

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

Home > cosmosdb14

cosmosdb14 | Data Explorer

Azure Cosmos DB account

Search

Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Cost Management
Quick start
Notifications
Data Explorer
Settings
Features
Replicate data globally
Default consistency
Backup & Restore
Networking
CORS

NOSQL API

DATA

- ToDoList14
 - Scale
 - id14
 - Items
 - Settings
 - Stored Procedures
 - User Defined Functions
 - Triggers

NOTEBOOKS

Notebooks is currently not available. We are working on it.

Home id14 - Items x Settings

SELECT * FROM c

Edit Filter

id /id

Load more

```
1 {
2   "id": "1",
3   "category": "student",
4   "Name": "Harshada",
5   "subject": "Data on cloud practical",
6   "iscomplete": false
7 }
```

cosmosdb14 - Microsoft Azure x New Tab x +

portal.azure.com/7Microsoft_Azure_Education_correlationId=1e92a84dc9e540a5b8c3c89cceab77a2&Microsoft_Azure_Education_newA4E=true&Microso...

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

Home > cosmosdb14

cosmosdb14 | Data Explorer

Azure Cosmos DB account

Search

Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Cost Management
Quick start
Notifications
Data Explorer
Settings
Features
Replicate data globally
Default consistency
Backup & Restore
Networking
CORS

NOSQL API

DATA

- ToDoList14
 - Scale
 - id14
 - Items
 - Settings
 - Stored Procedures
 - User Defined Functions
 - Triggers

NOTEBOOKS

Notebooks is currently not available. We are working on it.

Home id14 - Items x Settings Query 1

SELECT * FROM c

Edit Filter

id /id

Load more

```
1 {
2   "id": "1",
3   "category": "student",
4   "Name": "Harshada",
5   "subject": "Data on cloud practical",
6   "iscomplete": false,
7   "_rid": "q0u0ANoi3ZYBAAAAAAAAA==",
8   "_self": "dbs/q0u0AA==/colls/q0u0ANoi3ZY+/-docs/q0u0ANoi3ZYBAAAA",
9   "_etag": "\"2e805f2f-0000-0700-0000-038cb1600000\"",
10  "_attachments": "attachments/",
11  "_ts": 1670164845
12 }
```

Successfully created new item for container id14

cosmosdb14 - Microsoft Azure x New Tab x +

portal.azure.com/?Microsoft_Azure_Education_correlationId=1e92a84dc9e540a5b8c3c89cceb77a2&Microsoft_Azure_Education_newA4E=true&Microso...

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

Home > cosmosdb14

cosmosdb14 | Data Explorer

Azure Cosmos DB account

Search

Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Cost Management
Quick start
Notifications
Data Explorer
Settings
Features
Replicate data globally
Default consistency
Backup & Restore
Networking
CORS

NOSQL API

DATA

- ToDoList14
 - Scale
 - id14
 - Items
 - Settings
 - Stored Procedures
 - User Defined Functions
 - Triggers

NOTEBOOKS

Notebooks is currently not available. We are working on it.

Home | id14 - Items | Settings | Query 1 x

1 SELECT * FROM c

Results Query Stats

1 - 1

```
{
  "id": "1",
  "category": "student",
  "Name": "Harshada",
  "subject": "Data on cloud practical",
  "iscomplete": false,
  "_rid": "BFUXAOvckVUCAAAAAAAAAA==",
  "_self": "dbs/BFUXAA==/colls/BFUXAOvckVU=/docs/BFUXAOvckVU=/documents/1",
  "_etag": "\"d60030cd-0000-0700-0000-638cd7f30000\"",
  "_attachments": "attachments/",
  "_ts": 1670174707
}
```

Home | Settings | id14 - Items x | Query 1

SELECT * FROM c Edit Filter

id /id

Load more

```
1 {
2   "id": "1",
3   "category": "student",
4   "Name": "Harshada",
5   "subject": "Data on cloud practical",
6   "iscomplete": false,
7   "_rid": "BFUXAOvckVUCAAAAAAAAAA==",
8   "_self": "dbs/BFUXAA==/colls/BFUXAOvckVU=/docs/BFUXAOvckVU=/documents/1",
9   "_etag": "\"d60030cd-0000-0700-0000-638cd7f30000\"",
10  "_attachments": "attachments/",
11  "_ts": 1670174707
12 }
```

0 x 1 i 10