

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

Aim : Suppose you are a data engineer working on a project that involves extracting data from IBM Cloud Object Storage and loading it into a DB2 database. The data in the Object Storage consists of CSV files containing sales records for various products. Your task is to load the data to DB2 and perform SQL queries on the loaded data.

Data Extraction and Loading :

Describe the steps you would take to extract data from IBM Cloud Object Storage.

Explain how you would establish a connection to the Object Storage service and retrieve the CSV files containing pizza orders records.

Outline the process of loading these CSV files into a DB2 database table.

Database Table Design :

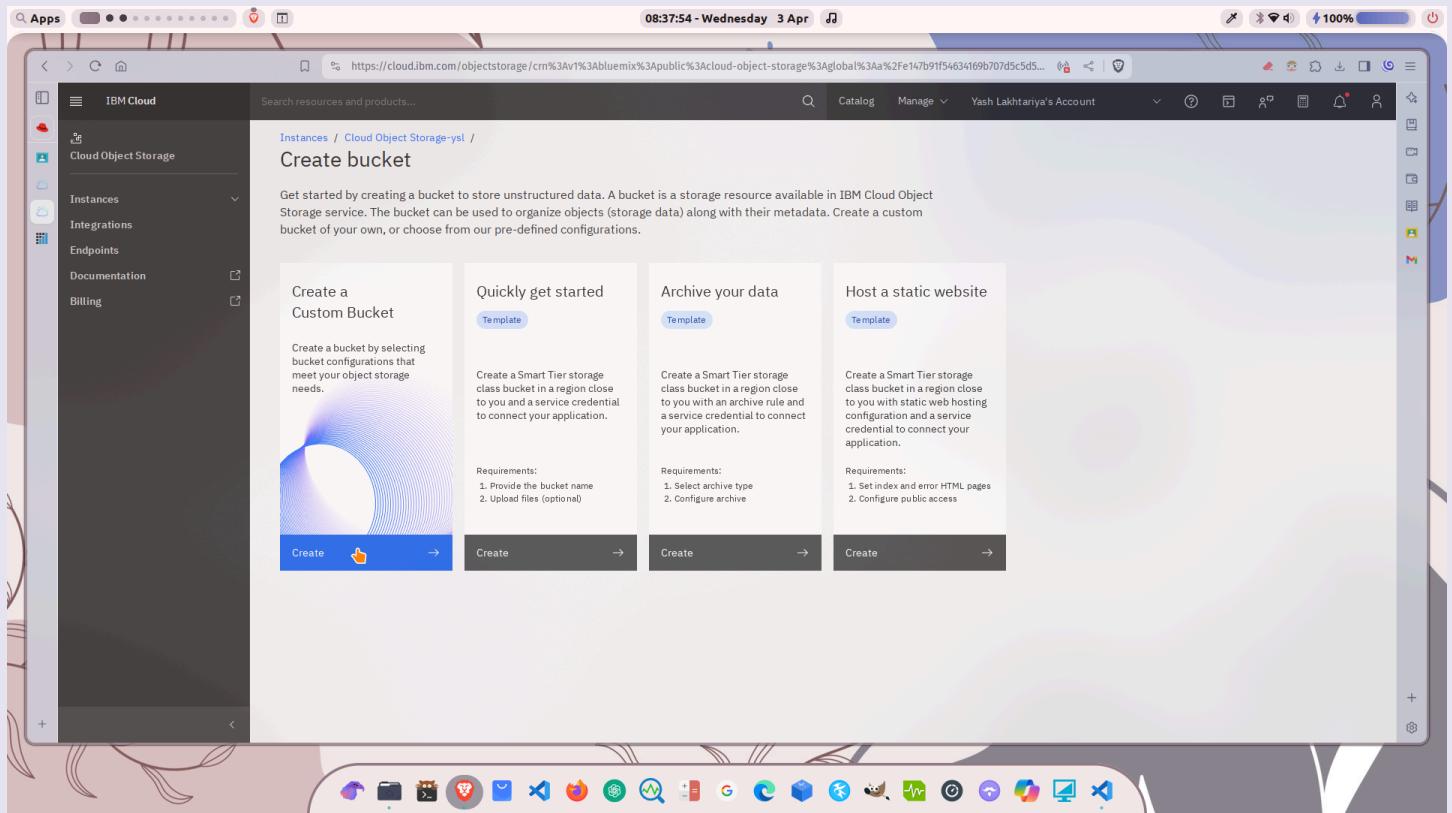
Design a DB2 database table schema to store the orders records. Consider the structure of the CSV files and determine appropriate data types for each column.

SQL Queries and Analysis :

Provide SQL queries to perform the analyses on the loaded orders data

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

1. Create a custom bucket in Cloud Object Storage by IBM

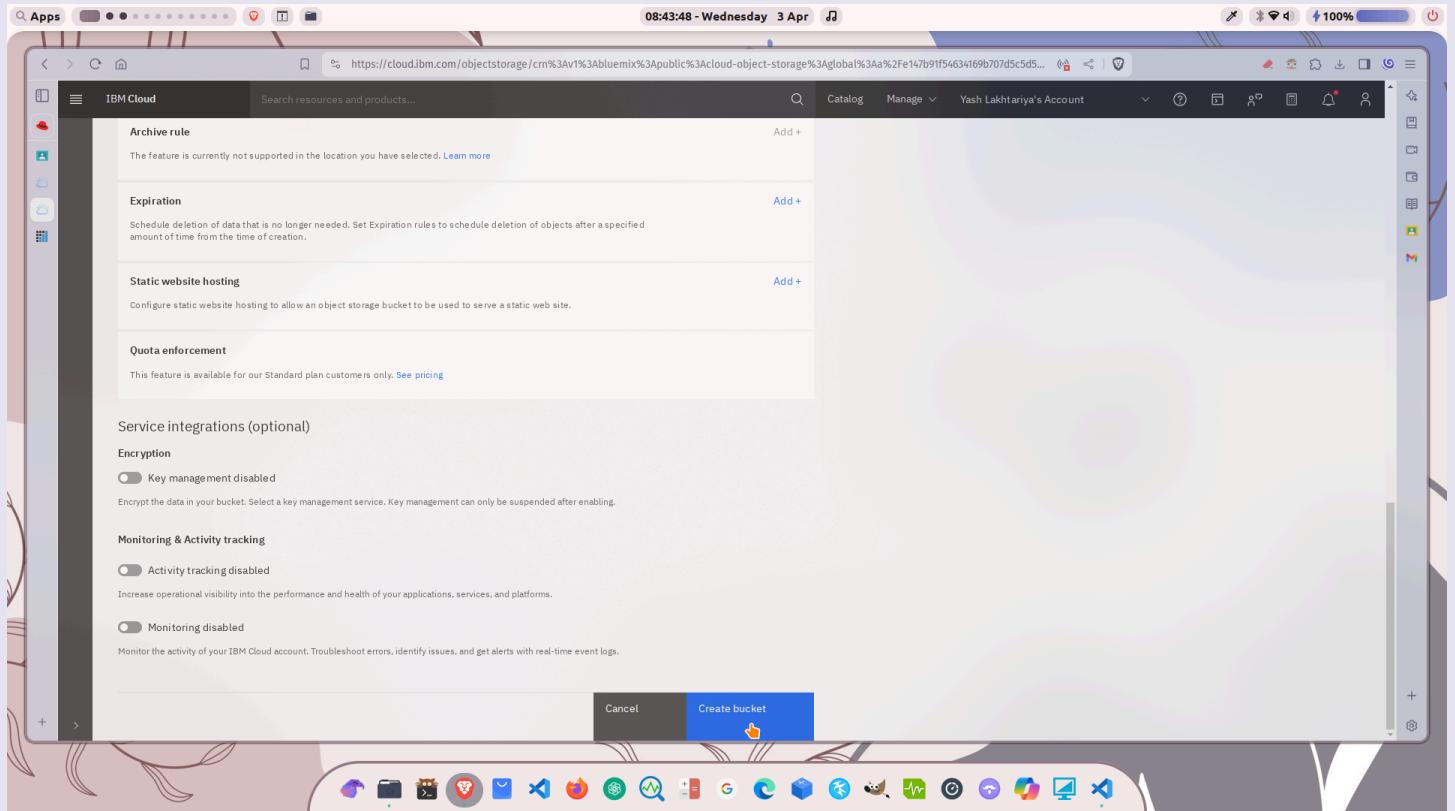


The screenshot shows the 'Create bucket' page in the IBM Cloud Object Storage interface. The URL in the address bar is <https://cloud.ibm.com/objectstorage/crn%3Av1%3Abuemix%3Apublic%3Acloud-object-storage%3Aglobal%3Aa%2Fe147b91f54634169b707d5c5d5...>. The page title is 'Create bucket'. On the left, there's a sidebar with 'IBM Cloud' and 'Cloud Object Storage' selected. The main content area has a heading 'Get started by creating a bucket to store unstructured data. A bucket is a storage resource available in IBM Cloud Object Storage service. The bucket can be used to organize objects (storage data) along with their metadata. Create a custom bucket of your own, or choose from our pre-defined configurations.' Below this, there are four options:

- Create a Custom Bucket**: 'Create a bucket by selecting bucket configurations that meet your object storage needs.' It includes a large blue button labeled 'Create' with a hand cursor icon.
- Quickly get started**: 'Create a Smart Tier storage class bucket in a region close to you and a service credential to connect your application.' It includes a 'Template' button and a 'Create' button.
- Archive your data**: 'Create a Smart Tier storage class bucket in a region close to you with an archive rule and a service credential to connect your application.' It includes a 'Template' button and a 'Create' button.
- Host a static website**: 'Create a Smart Tier storage class bucket in a region close to you with static web hosting configuration and a service credential to connect your application.' It includes a 'Template' button and a 'Create' button.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

2. Assign the name and select cross region



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

3. Check its configuration and proceed further

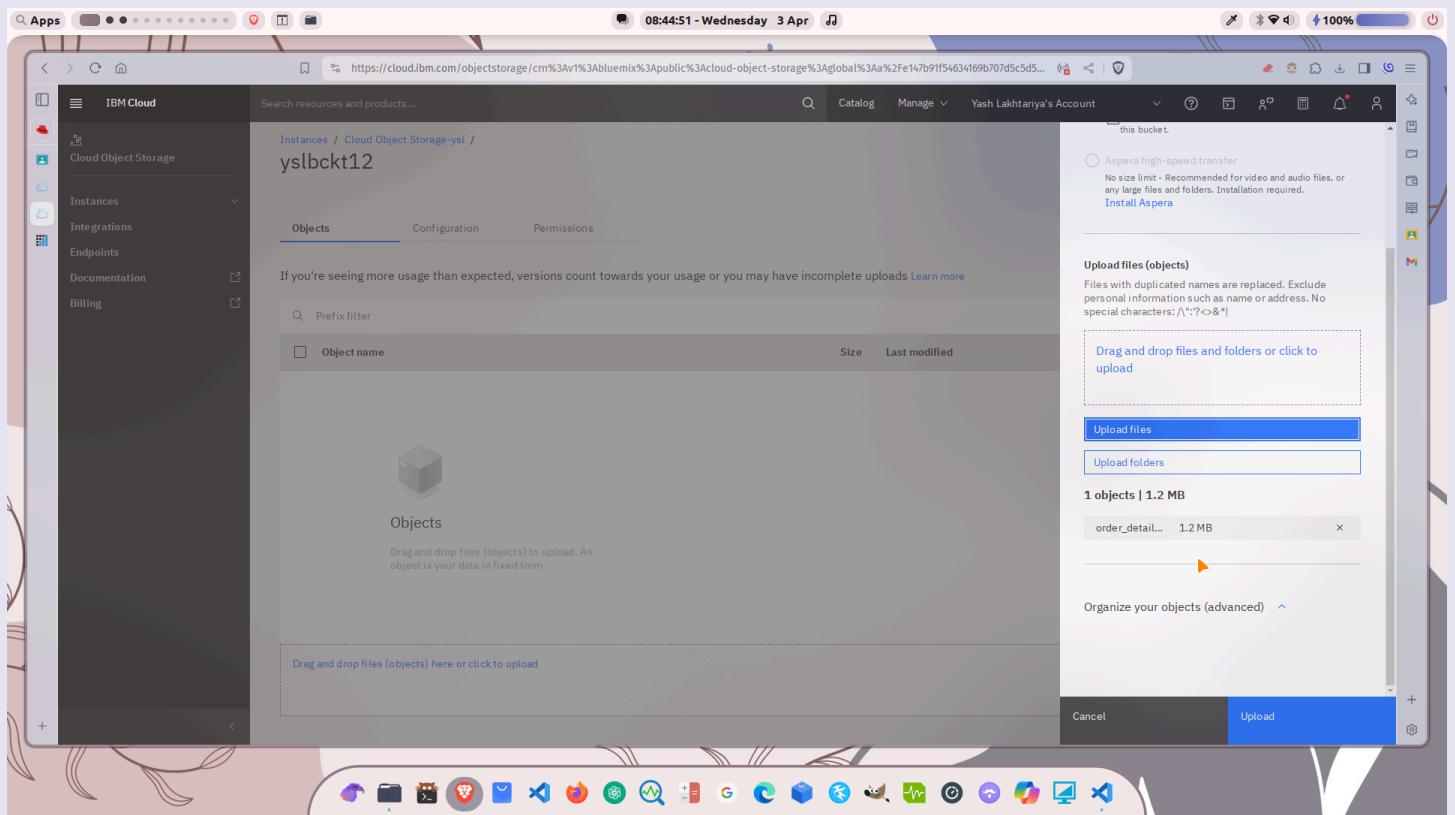
The screenshot shows the IBM Cloud Object Storage configuration page for a bucket named 'yslbckt12'. The left sidebar includes 'IBM Cloud', 'Cloud Object Storage', 'Instances', 'Integrations', 'Endpoints', 'Documentation', and 'Billing'. The main content area has tabs for 'Objects', 'Configuration', and 'Permissions', with 'Configuration' selected. Under 'Bucket configuration', there's a 'Jump to' section with links for 'Bucket details', 'Endpoints', 'Versioning', 'Object Lock', 'Key management', 'Activity Tracker', 'Monitoring', and 'Lifecycle policy'. The 'Bucket details' section shows the following information:

Bucket name	yslbckt12	Total bytes	0 bytes
Service instance	cloud-object-storage	Resiliency	Cross Region
Total objects	0	Location	United States Geo (us-geo)
Storage class	Smart Tier	Date created	2024-04-03 8:43 AM
Cloud Functions trigger	Disabled	Learn more	

The 'Endpoints' section contains a CRN (Cloud Resource Name) for the service instance: `crn:v1:bluemix:public:cloud-object-storage:global:a/e147b91f54634169b707d5c5d51b9c77:c2b0aeef9-875f-4e7c-a336-4397d8190135:bucket:yslbckt12`. Below this, a note states: "Endpoints are used hand in hand with your credentials (i.e. keys, CRN, bucket name) to tell your service where to look for this bucket. Depending on where your service or application is located you will want to use one of the below endpoint types." A row of icons for various cloud services like AWS, Google Cloud, and Microsoft Azure is shown at the bottom.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

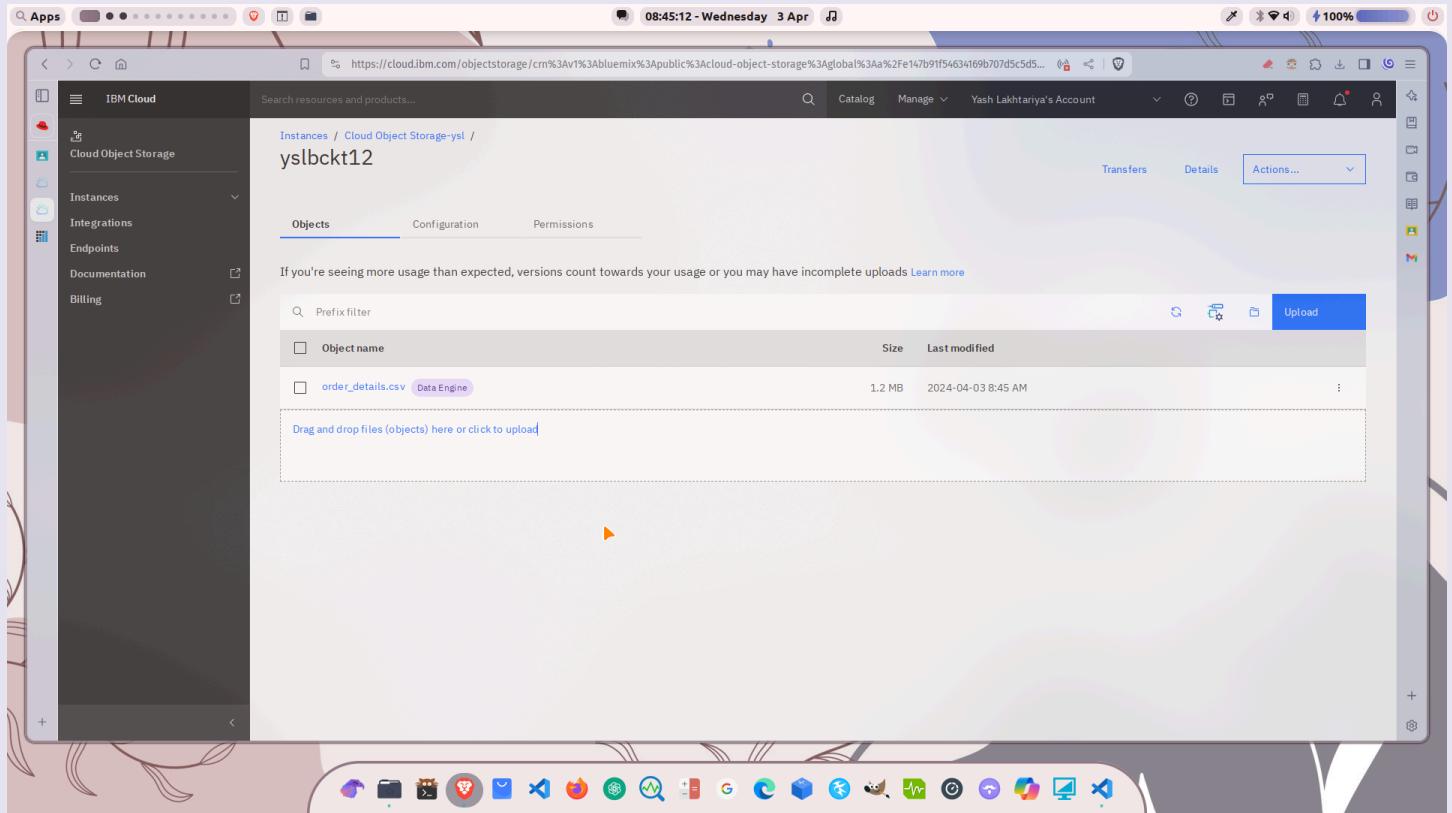
4. In the bucket tab, select upload and upload files option and upload csv file



The screenshot shows the IBM Cloud Object Storage interface. On the left, a sidebar menu includes 'Cloud Object Storage' under 'Instances'. The main area displays a bucket named 'yslbckt12'. A sub-menu bar at the top of the main area shows 'Objects', 'Configuration', and 'Permissions'. Below this, there's a message about usage and incomplete uploads. A search bar and a 'Prefix filter' input field are present. The central part of the screen shows a large 'Objects' section with a placeholder message: 'Drag and drop files (objects) to upload. An object is your data in fixed form.' Two upload areas are visible: one on the left labeled 'Drag and drop files (objects) here or click to upload' and another on the right labeled 'Upload files (objects)' with a sub-section for 'Upload files' and 'Upload folders'. A progress bar indicates '1 objects | 1.2 MB' for a file named 'order_detail...'. At the bottom, there are 'Cancel' and 'Upload' buttons.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

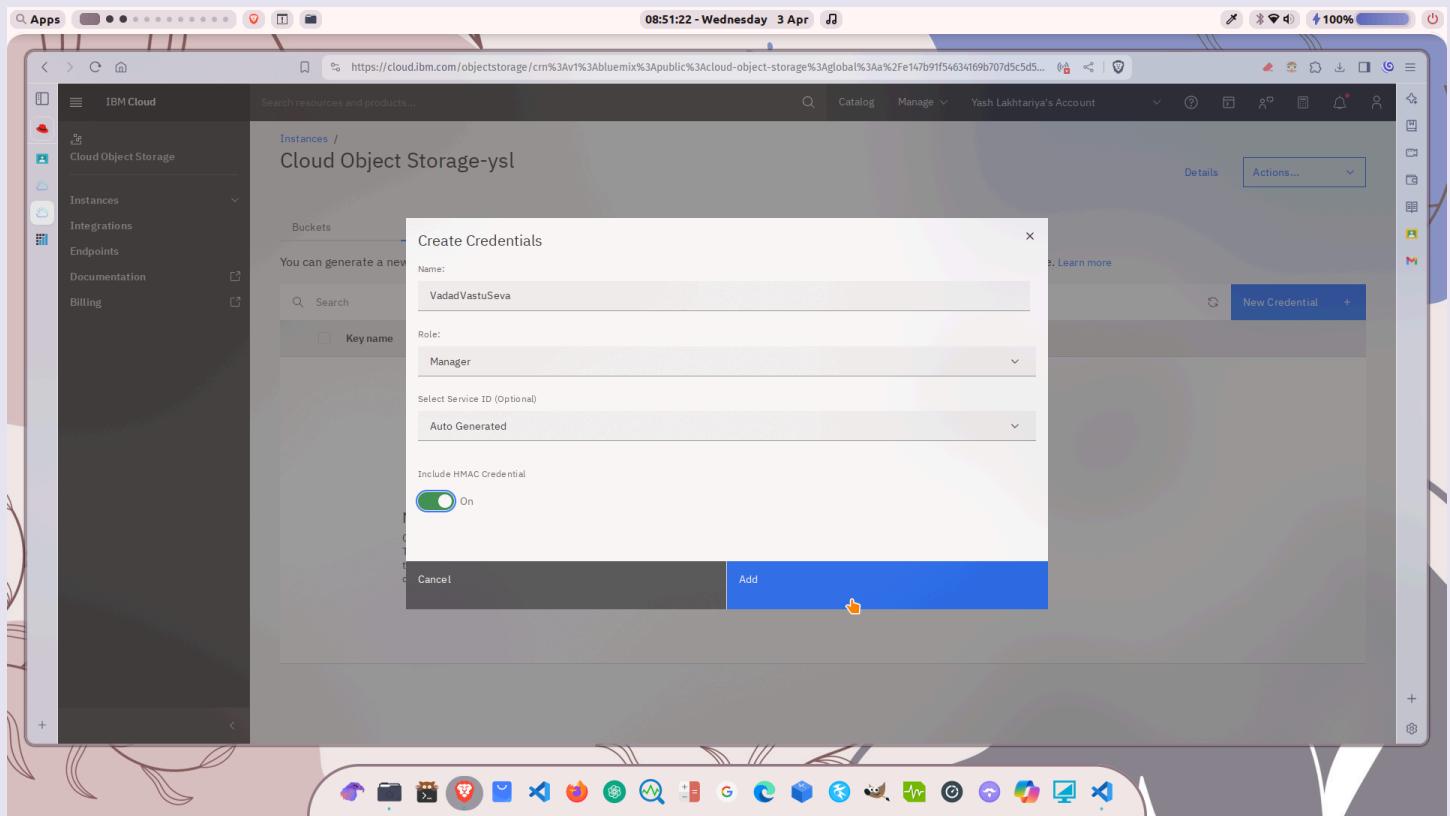
5. The file order.csv is uploaded in the bucket



The screenshot shows the IBM Cloud Object Storage interface. The left sidebar has 'Cloud Object Storage' selected. The main area shows a bucket named 'yslbckt12'. Under the 'Objects' tab, there is one file listed: 'order_details.csv' (Data Engine). The file was uploaded on April 3, 2024, at 8:45 AM, with a size of 1.2 MB. A large orange arrow points downwards towards the upload button.

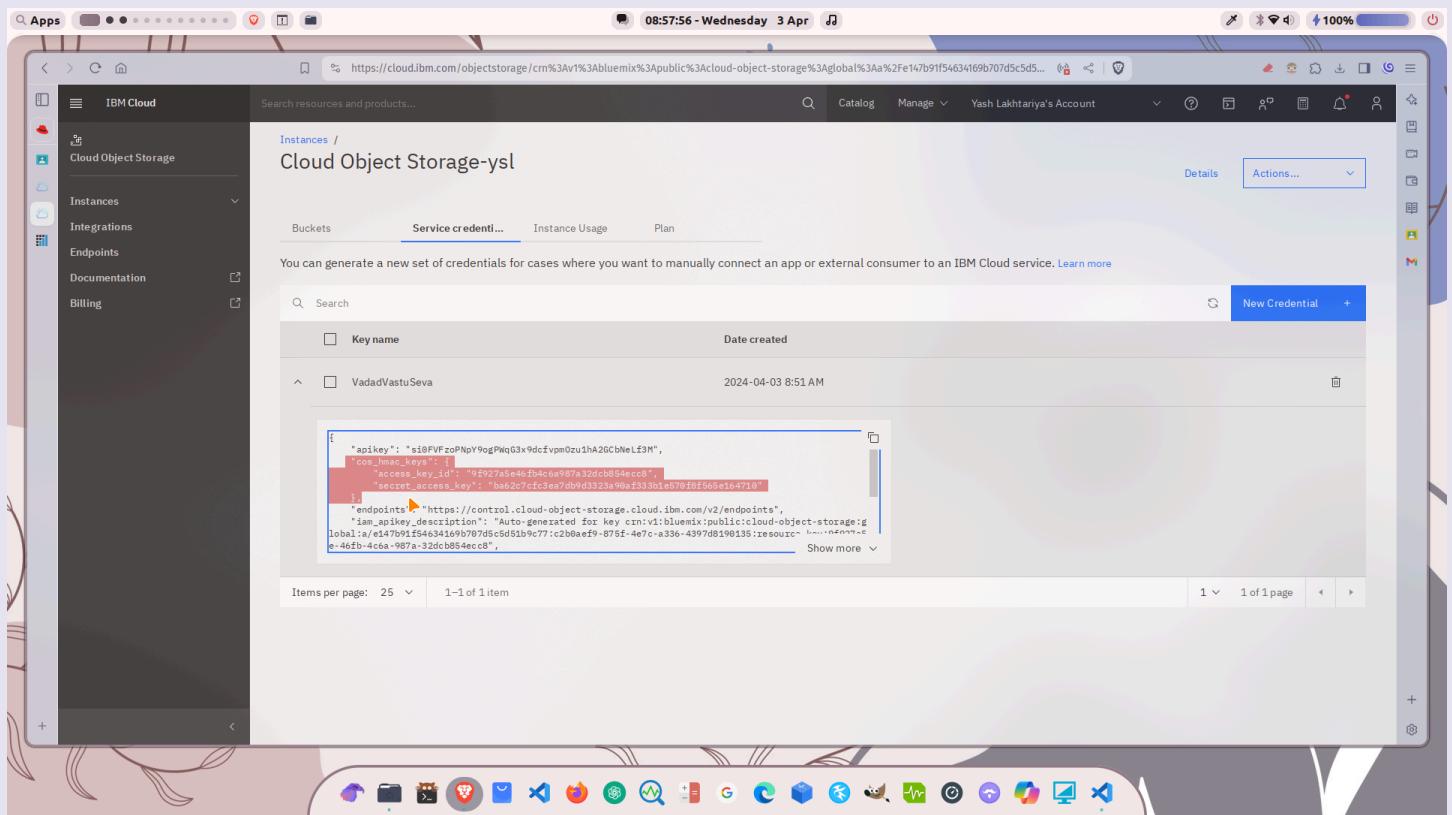
Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

6. Now, in Object storage's Service Credentials tab, create service credentials or Manager Role



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

7. From service credentials, copy the cos_hmac_keys pair



The screenshot shows the IBM Cloud Object Storage service credentials page. The URL is https://cloud.ibm.com/objectstorage/crn%3Av1%3Abuemix%3Apublic%3Acloud-object-storage%3Aglobal%3Aa%2Fe147b91f54634169b707d5c5d5... The page displays a table of credentials, with one row selected. The selected row contains the following JSON data:

```
[{"apikey": "si@FVFzoPNpY9ogPWqG3x9dcfvpm0zu1hA2GcbNeLz3M", "cos_hmac_keys": [{"access_key_id": "9f927a8e46fb4c6987a32dcba54ecc8", "secret_access_key": "ba62c7fcf3e709d3323a98af33351e579fdff965e164710"}, {"endpoints": "https://control.cloud-object-storage.cloud.ibm.com/v2/endpoints", "iam_apikey_description": "Auto-generated for key crn:v1:bluemix:public:cloud-object-storage:global:a/e147b91f54634169b707d5d5b19c77:c2b0ae9-876f-4e7c-a336-4397d8190135:resource-e-46fb-4cfa-987a-32dcbb854ecc8"}]
```

Name - Yash Lakhtariya

Enrollment number - 21162101012

Branch - CBA Batch - 61

EADC Practical 12

8. Visit Db2 instance created earlier (create if not exists) and in Data tab, load data from Cloud Object storage, paste the credentials, access key and secret access key there

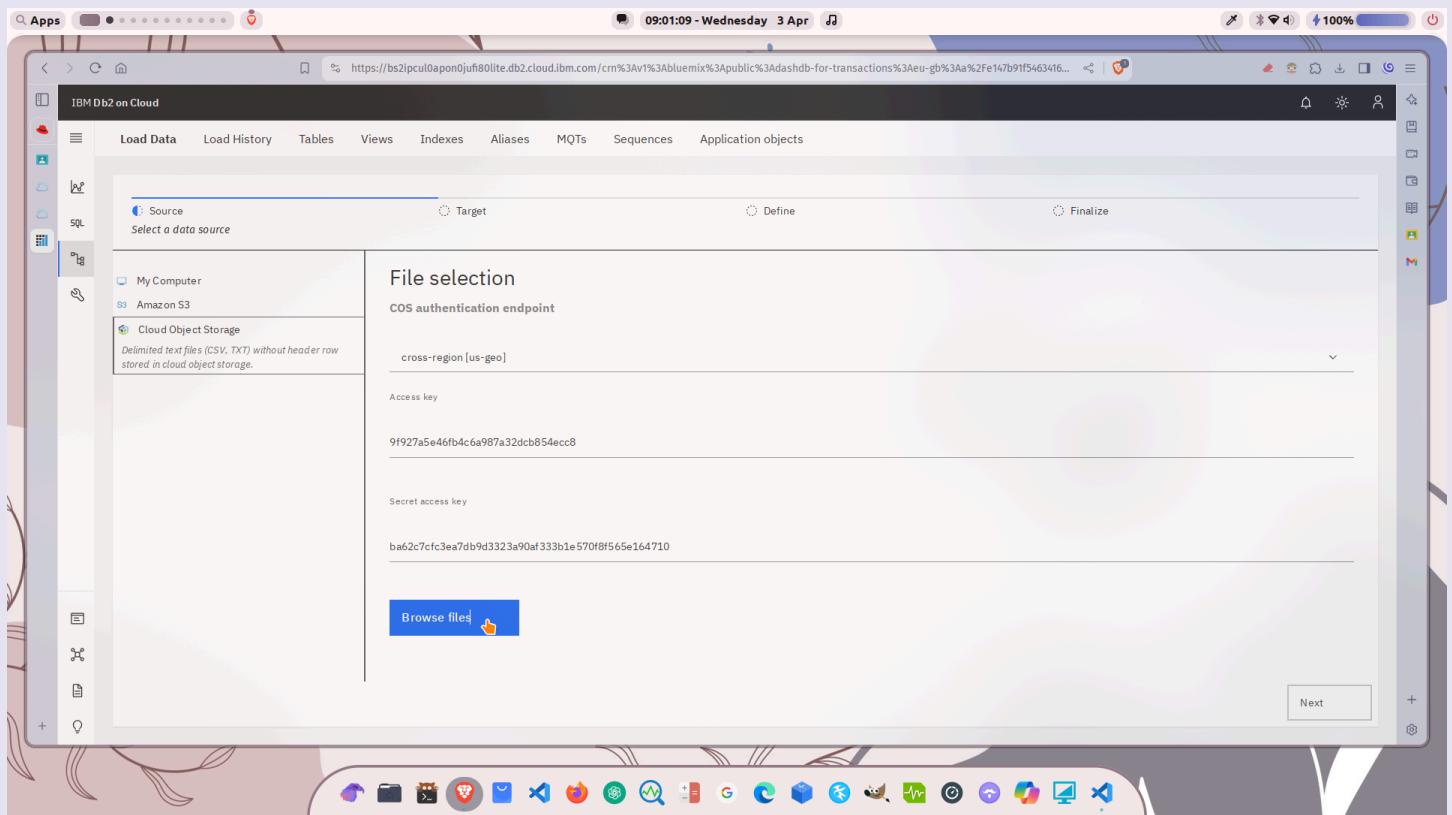
The screenshot shows the IBM Db2 on Cloud interface. In the top navigation bar, the 'Data' tab is selected. On the left sidebar, the 'Load Data' icon is highlighted. The main panel shows the 'Source' tab selected, with a dropdown menu showing 'Cloud Object Storage' is chosen. A 'File selection' dialog is open, displaying the following information:

- COS authentication endpoint: cross-region [us-geo]
- Access key: 9f927a5e46fb4c6a987a32dcb854ecc8
- Secret access key: ba62c7fc3ea7db9d3323a90af333b1e570f8f565e164710

A blue 'Browse files' button is visible at the bottom of the dialog. In the bottom right corner of the dialog, there is a 'Next' button.

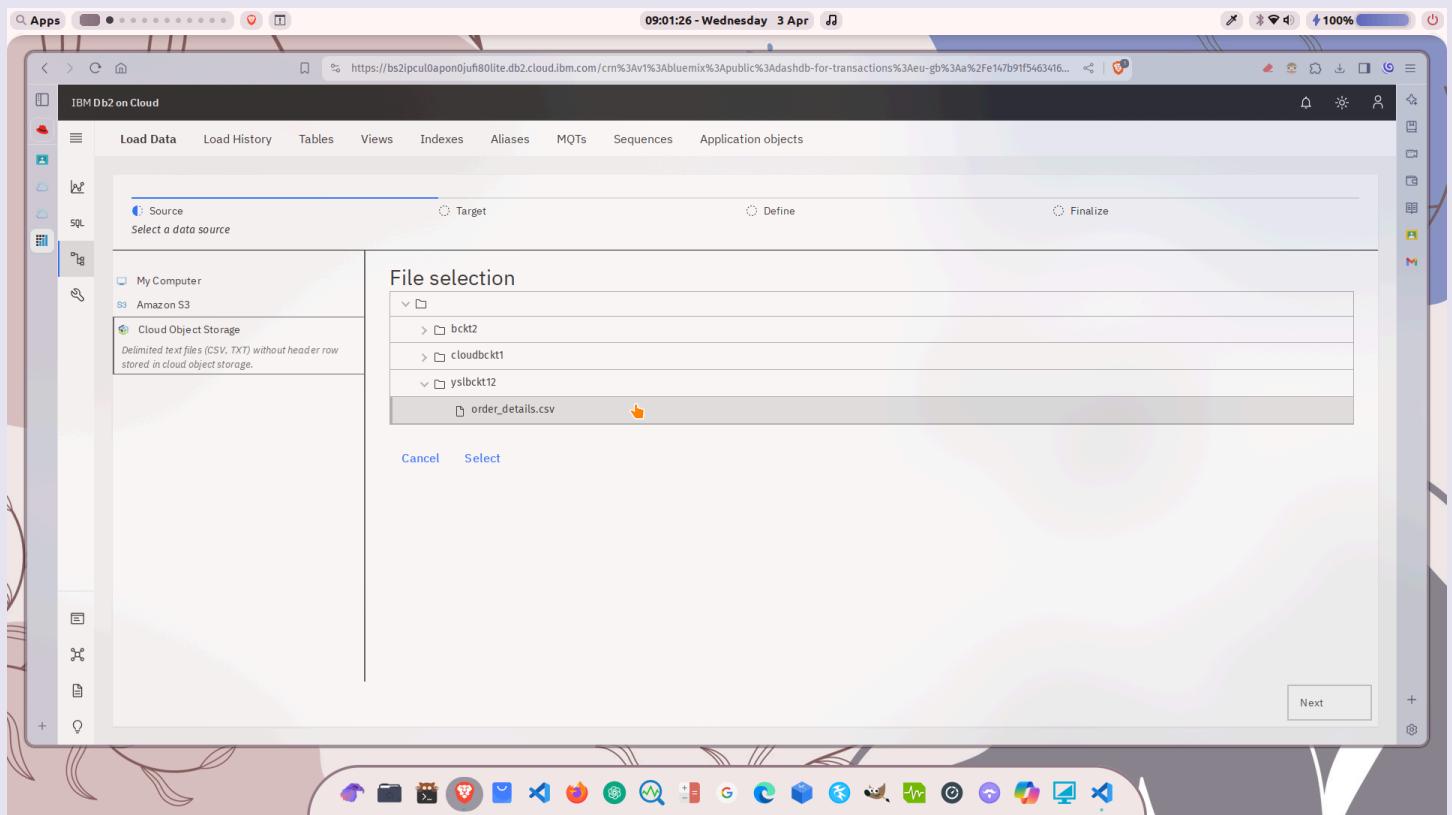
Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

9. Click Browse files option to get our file loaded from object storage



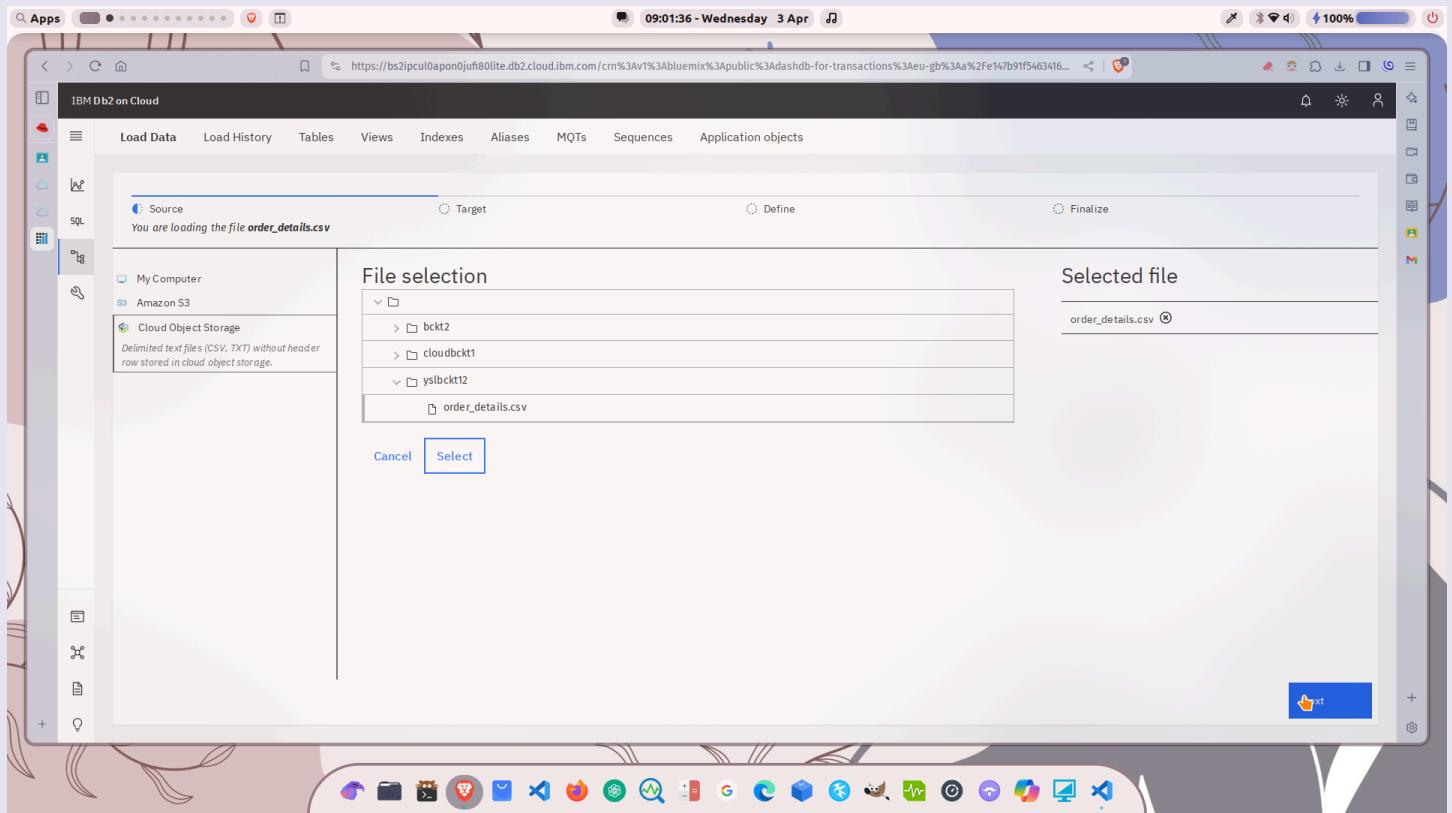
Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

10. Find the bucket and browse the required csv file



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

11. Select the file and click next



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

12. Now, in target selection create table from the csv file for Db2 database

The screenshot shows the IBM Db2 on Cloud interface. The top navigation bar includes 'Load Data', 'Load History', 'Tables', 'Views', 'Indexes', 'Aliases', 'MQTS', 'Sequences', and 'Application objects'. The 'Load Data' tab is active. The main area is titled 'Select a load target' and shows 'Schema' (WQY86844) and 'Table' (EMPLOYEES). To the right, there's a panel for defining a new table named 'order' with the following columns:

```
ORDERDTLSID INT
ORDERID INT
PIZZAID VARCHAR(100)
QUANTITY INT
```

A 'Create' button is present at the bottom of this panel. Navigation buttons 'Back' and 'Next' are also visible.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

13. Create and click next overriding the previous data

The screenshot shows the IBM Db2 on Cloud interface for loading data. The 'Source' tab is selected, indicating that 'order_details.csv' is being loaded into the 'WQY86844.ORDER' table. The target schema is 'WQY86844'. The 'Table' section lists 'EMPLOYEES' and 'ORDER'. On the right, the 'Table definition' for the 'ORDER' table is displayed, showing four columns: ORDERTLSID (INTEGER), ORDERID (INTEGER), PIZZIID (VARCHAR), and QUANTITY (INTEGER). The 'NULLABLE' column has 'Y' for all columns. An option to 'Overwrite table with new data' is selected. At the bottom right, there are 'Back' and 'Next' buttons, with the 'Next' button being highlighted by a mouse cursor.

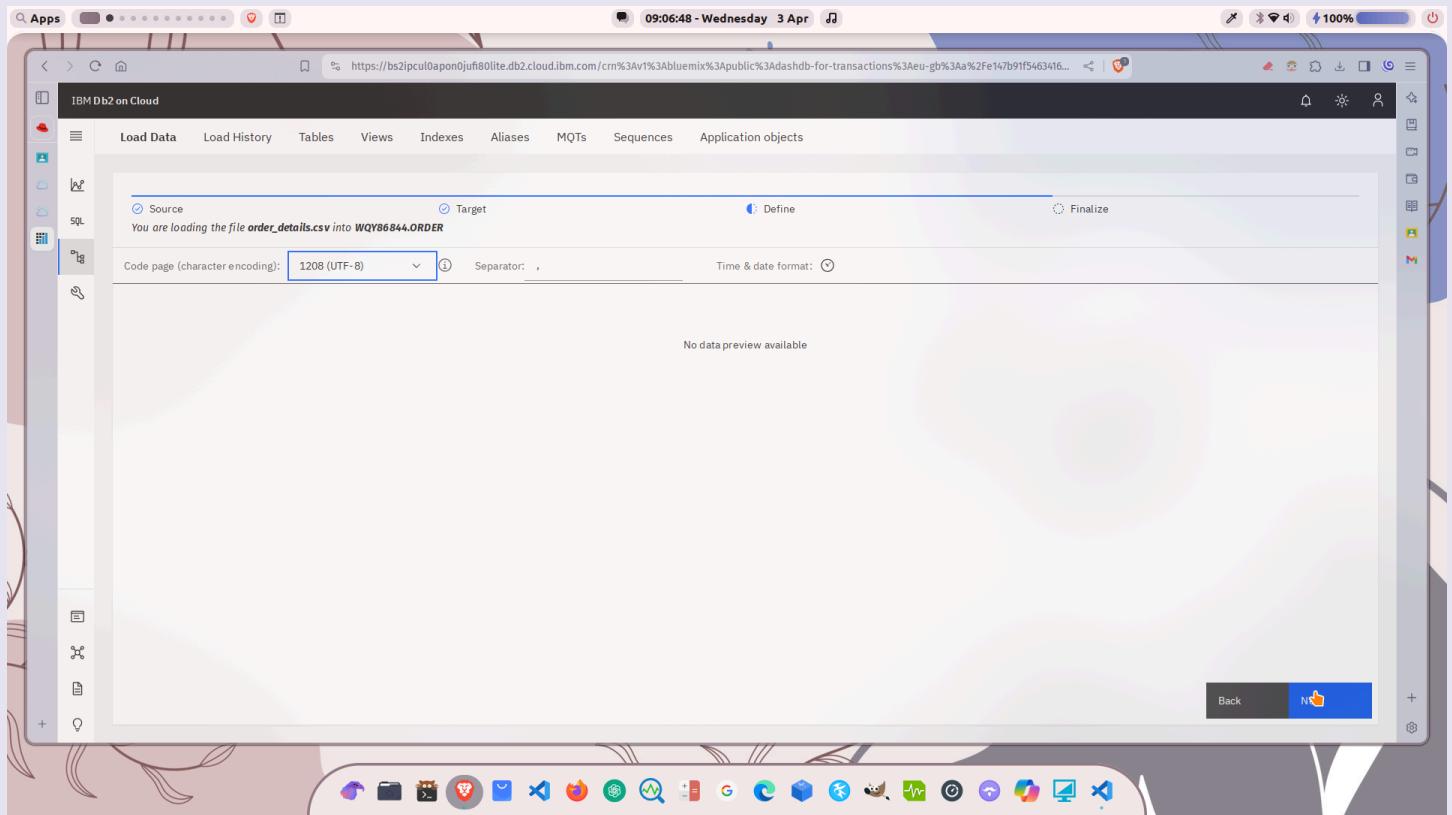
Name - Yash Lakhtariya

Enrollment number - 21162101012

Branch - CBA Batch - 61

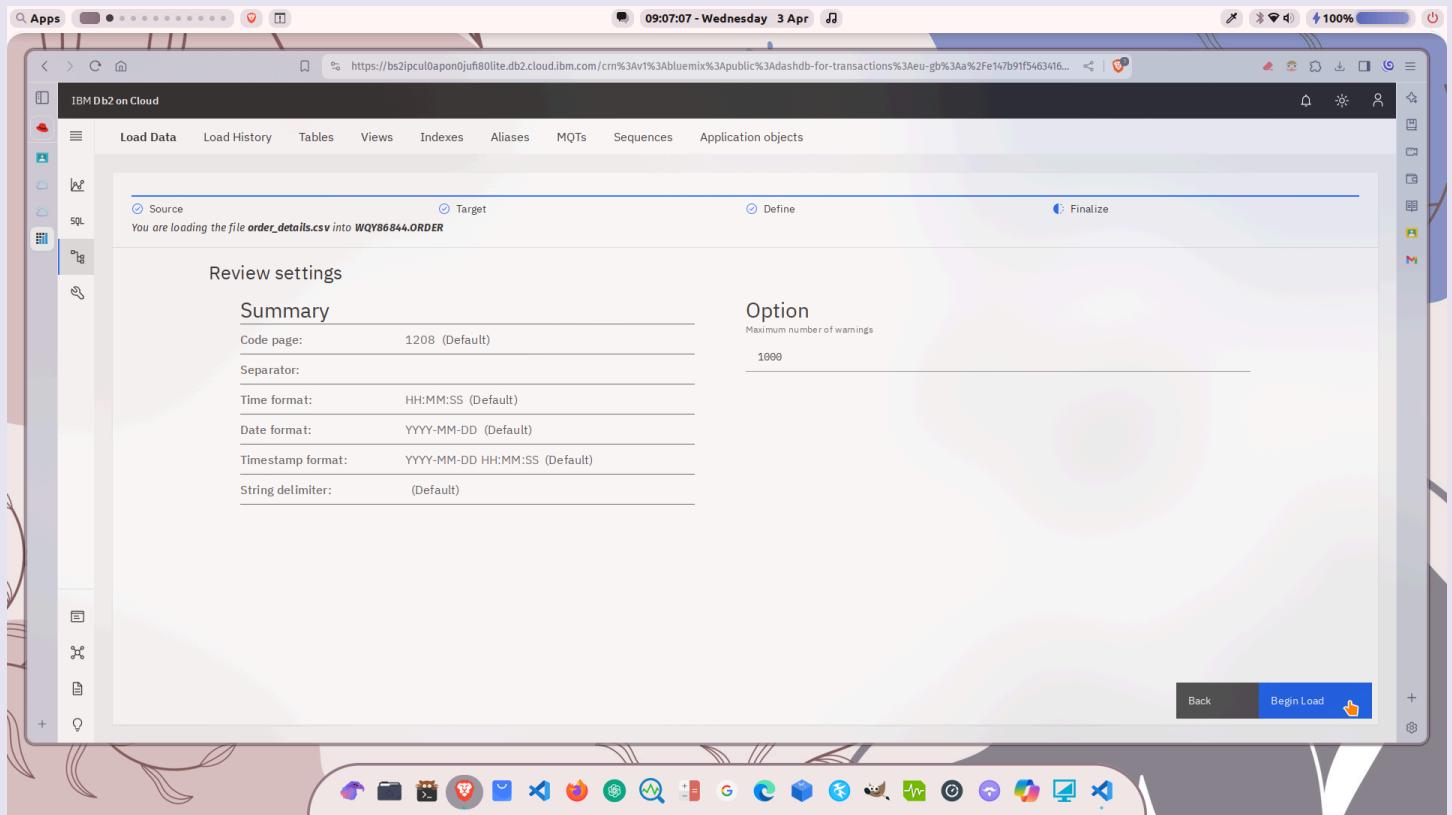
EADC Practical 12

14. Keep rest options like encoding and separator default and click next



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

15. Check the configuration and begin loading the data



The screenshot shows the IBM Db2 on Cloud interface for managing database objects. The main menu bar includes 'Load Data', 'Load History', 'Tables', 'Views', 'Indexes', 'Aliases', 'MQTs', 'Sequences', and 'Application objects'. The 'Load Data' tab is selected.

The current screen displays the configuration for loading a file named 'order_details.csv' into a target table 'WQY86844.ORDER'. The 'Source' tab is selected, indicated by a blue outline. The target table is specified as 'WQY86844.ORDER'. There are four tabs at the top: 'Source' (selected), 'Target', 'Define', and 'Finalize'.

The 'Review settings' section contains two main sections: 'Summary' and 'Option'.

Summary:

- Code page: 1208 (Default)
- Separator: (empty)
- Time format: HH:MM:SS (Default)
- Date format: YYYY-MM-DD (Default)
- Timestamp format: YYYY-MM-DD HH:MM:SS (Default)
- String delimiter: (Default)

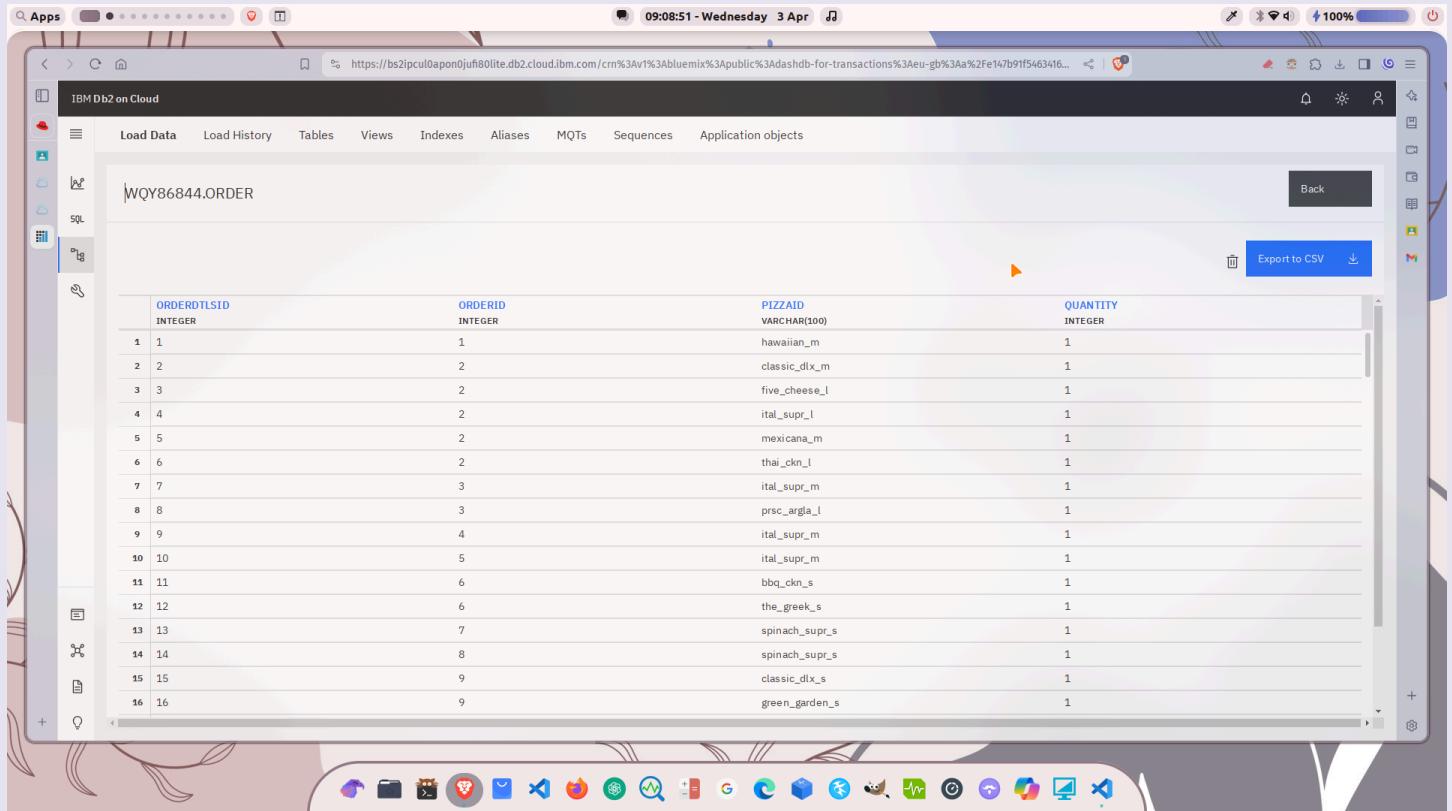
Option:

- Maximum number of warnings: 1000

At the bottom right of the interface, there are 'Back' and 'Begin Load' buttons. A hand cursor is hovering over the 'Begin Load' button.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

16. As seen, the data is successfully loaded

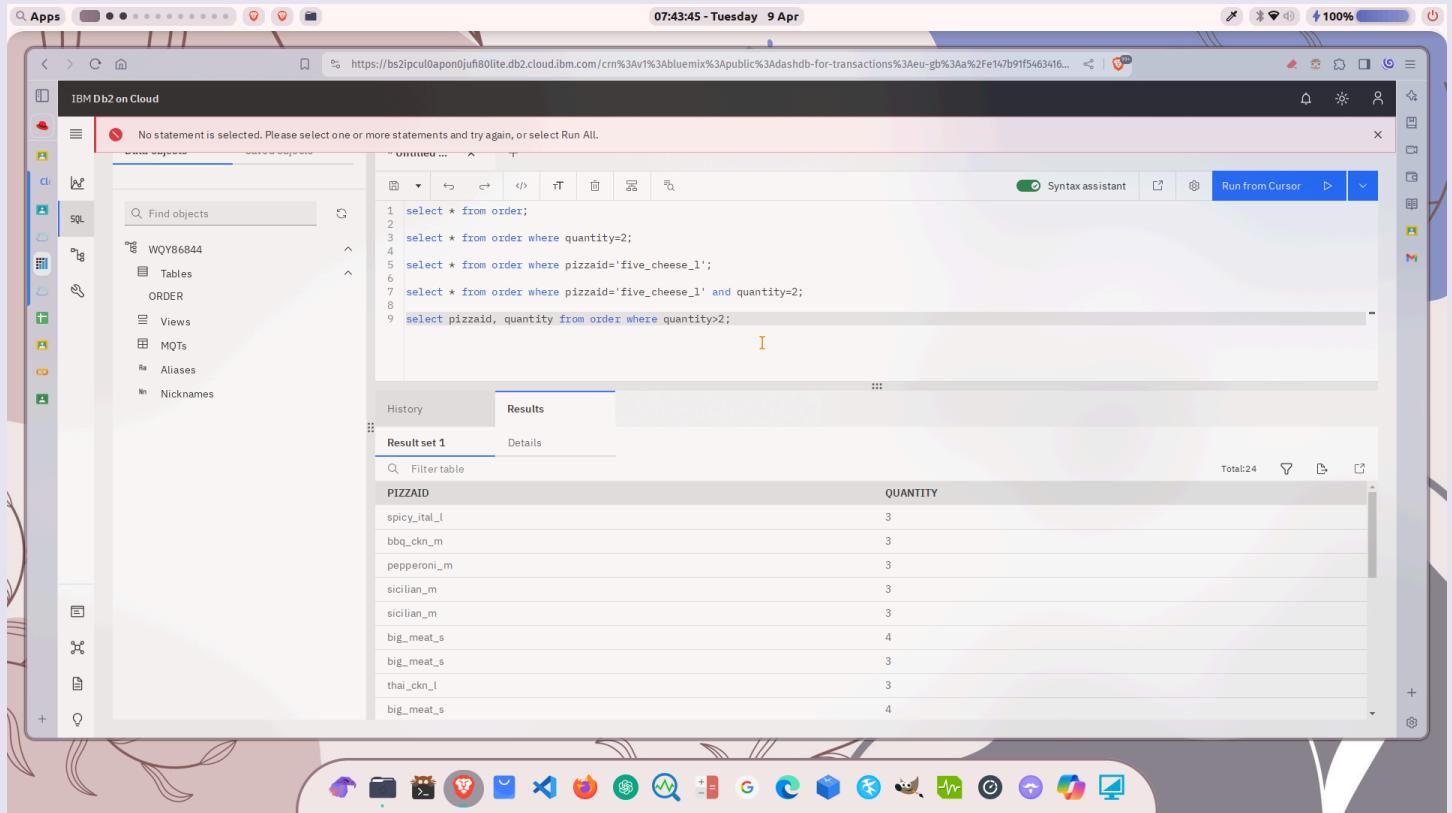


The screenshot shows the IBM Db2 on Cloud interface with the URL <https://bs2ipcul0apon0jufl80lite.db2.cloud.ibm.com/cm%3Av1%3Abluemix%3Apublic%3Adashdb-for-transactions%3Aeu-gb%3Aa%2Fe147b91f5463416...>. The page title is "IBM Db2 on Cloud". The main navigation bar includes "Load Data", "Load History", "Tables", "Views", "Indexes", "Aliases", "MQTs", "Sequences", and "Application objects". The current view is on the "Tables" tab, specifically the "WQY86844.ORDER" table. The table has four columns: ORDERDTLSID (INTEGER), ORDERID (INTEGER), PIZZAIID (VARCHAR(100)), and QUANTITY (INTEGER). The data consists of 16 rows, each with a unique ORDERDTLSID and ORDERID, and a corresponding PIZZAIID and QUANTITY value. A blue "Export to CSV" button is located in the top right corner of the table view.

	ORDERDTLSID INTEGER	ORDERID INTEGER	PIZZAIID VARCHAR(100)	QUANTITY INTEGER
1	1	1	hawaiian_m	1
2	2	2	classic_dlx_m	1
3	3	2	five_cheese_l	1
4	4	2	ital_supr_l	1
5	5	2	mexicana_m	1
6	6	2	thai_ckn_l	1
7	7	3	ital_supr_m	1
8	8	3	prsc_argla_l	1
9	9	4	ital_supr_m	1
10	10	5	ital_supr_m	1
11	11	6	bbq_ckn_s	1
12	12	6	the_greek_s	1
13	13	7	spinach_supr_s	1
14	14	8	spinach_supr_s	1
15	15	9	classic_dlx_s	1
16	16	9	green_garden_s	1

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

17. From Run SQL tab, run SQL queries on the order table



The screenshot shows the IBM Db2 on Cloud interface in a web browser. The URL is https://bs2ipcul0apon0jufl80lite.db2.cloud.ibm.com/cm%3Av1%3Abuemix%3Apublic%3dashdb-for-transactions%3Aeu-gb%3Aa%2Fe147b91f5463416... The top bar shows the date and time as 07:43:45 - Tuesday 9 Apr. The left sidebar has a 'SQL' icon selected, and the main area shows a list of statements:

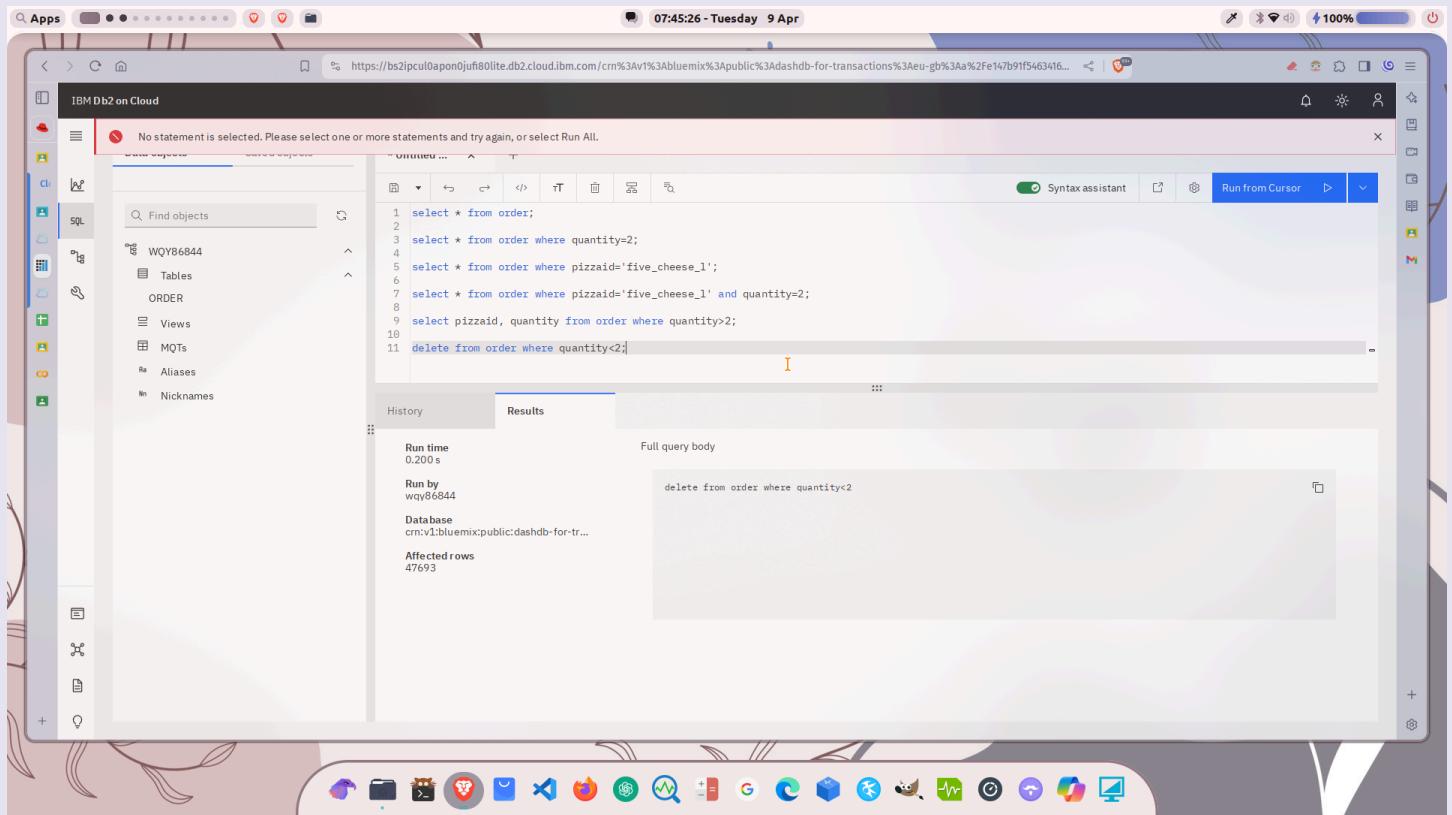
```
1 select * from order;
2
3 select * from order where quantity=2;
4
5 select * from order where pizzaaid='five_cheese_l';
6
7 select * from order where pizzaaid='five_cheese_l' and quantity=2;
8
9 select pizzaid, quantity from order where quantity>2;
```

The results pane shows a table with two columns: PIZZAID and QUANTITY. The data is as follows:

PIZZAID	QUANTITY
spicy_ital_l	3
bbq_ckn_m	3
pepperoni_m	3
sicilian_m	3
sicilian_m	3
big_meat_s	4
big_meat_s	3
thai_ckn_l	3
big_meat_s	4

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

18. Try queries which alter the table data also like delete



The screenshot shows the IBM Db2 on Cloud interface in a web browser. The URL is https://bs2ipcu0apon0jufl80lite.db2.cloud.ibm.com/cm%3Av1%3Abuemix%3Apublic%3Adashdb-for-transactions%3Aeu-gb%3Aa%2Fe147b91f5463416... The interface has a sidebar with various database objects like Tables, Views, MQTs, Aliases, and Nicknames. The main area is a SQL editor with the following code:

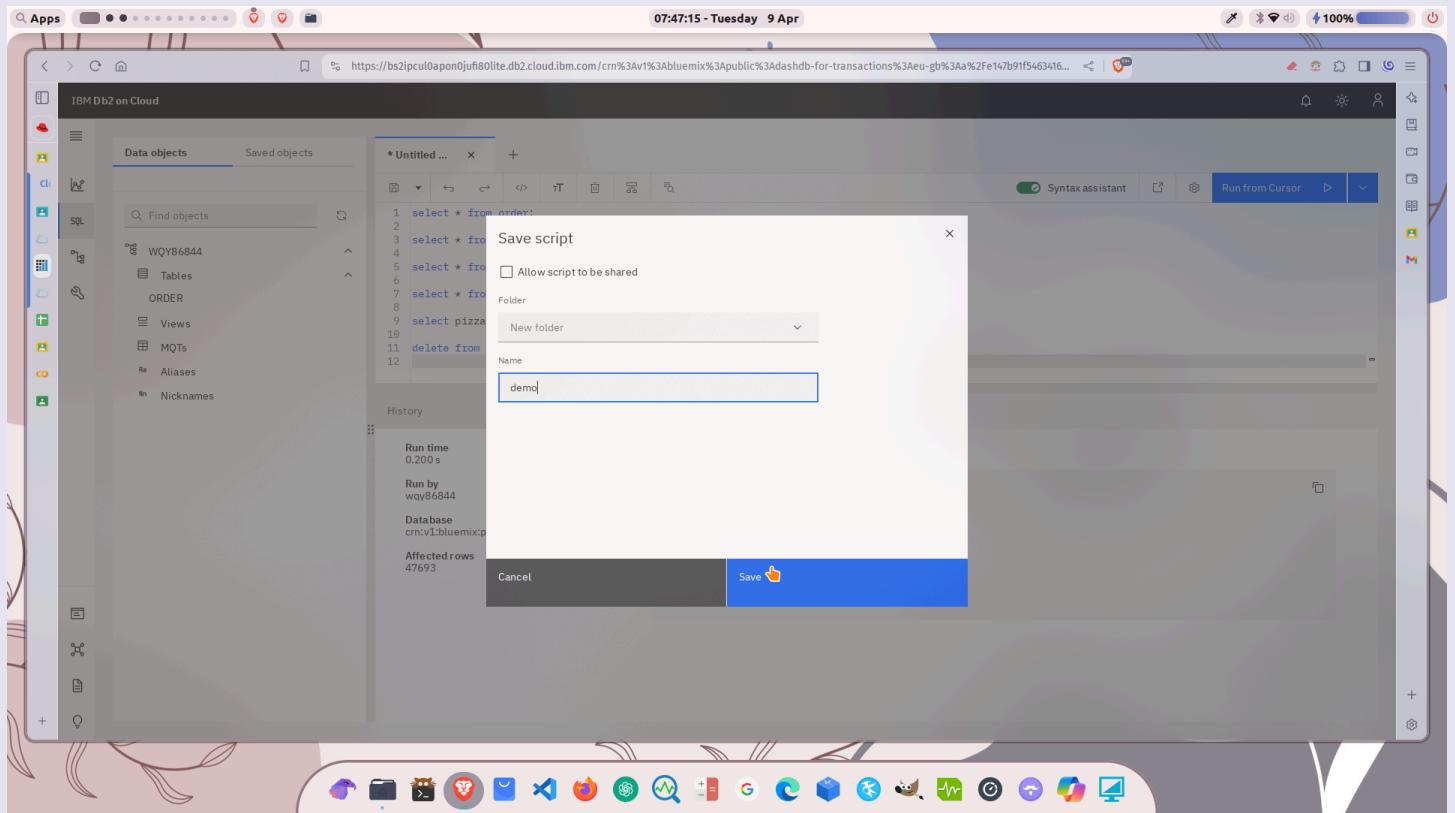
```
1 select * from order;
2
3 select * from order where quantity=2;
4
5 select * from order where pizzaid='five_cheese_1';
6
7 select * from order where pizzaid='five_cheese_1' and quantity=2;
8
9 select pizzaid, quantity from order where quantity>2;
10
11 delete from order where quantity<2;
```

Below the code, the results pane shows the following information:

- Run time: 0.200 s
- Run by: wqy86844
- Database: crn:v1:bluemix:public:dashdb-for-tr...
- Affected rows: 47693

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

19. SQL Script can also be saved to server with the option given



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

20. In options tab, the display properties can be changed for ease

The screenshot shows the IBM Db2 on Cloud interface in a web browser. The main window displays a SQL script named 'demo' containing several SELECT statements. A modal dialog titled 'Result panel options' is open over the script editor. The dialog contains the following settings:

- Maximum groups of results displayed (20-100): Set to 20.
- Maximum size of each result set returned (1-10000) KB: Set to 100.
- Auto-persist: An unchecked checkbox.
- Run results: An unchecked checkbox.
- Fetched result sets: An unchecked checkbox.
- Max number of result sets persisted (10-50): Set to 10.

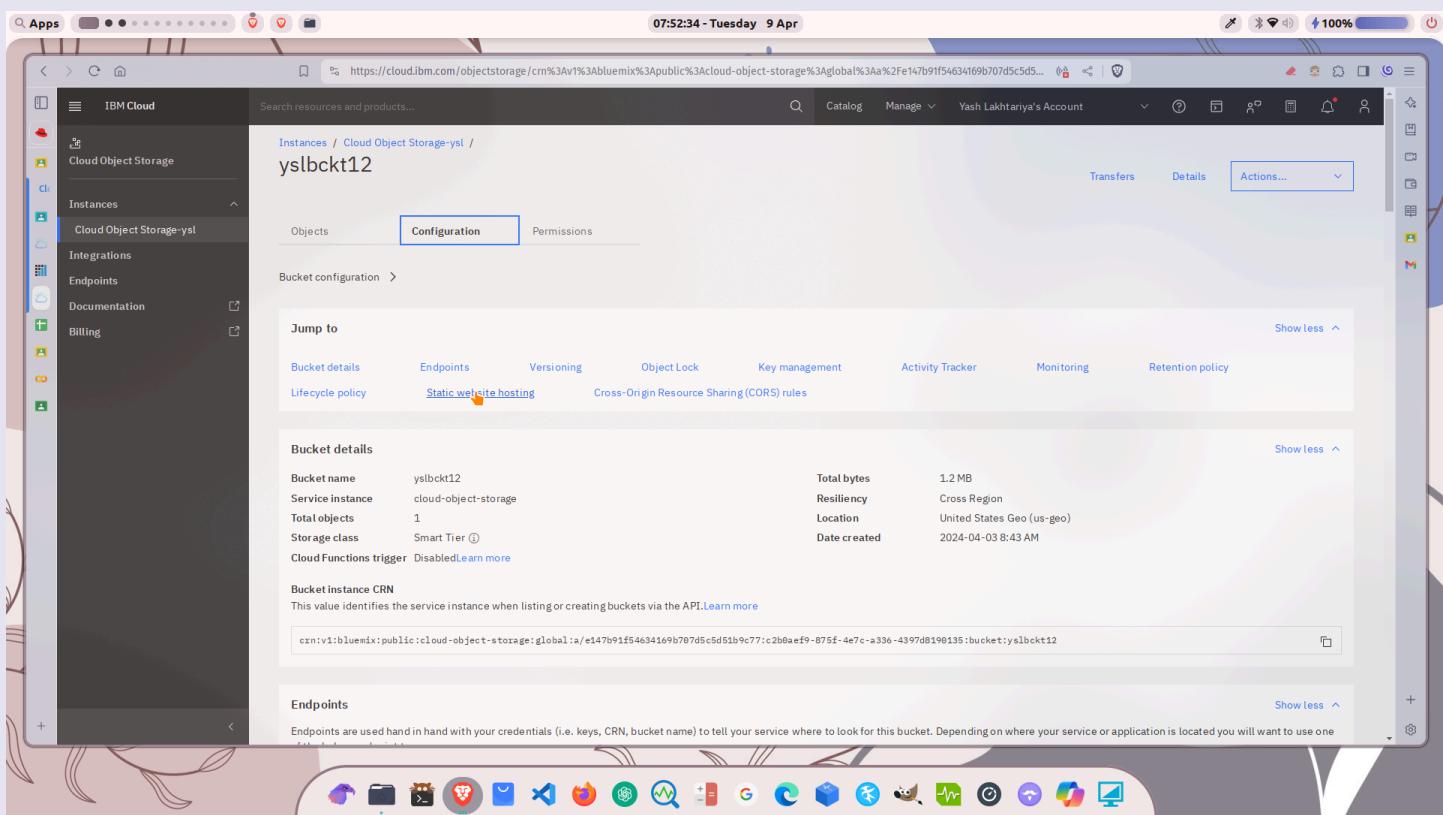
At the bottom of the dialog, there are 'Cancel' and 'Save' buttons. The background shows the results of the SQL queries, which are mostly empty or show small amounts of data. The status bar at the bottom indicates the date and time: Apr 9, 2024 7:43:14 AM.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

ADDITIONAL TASK :

Learn to host a static website on object storage and host your Portfolio or any sample static website of your own.

1. In bucket's configuration tab, click Static Website hosting option



The screenshot shows the IBM Cloud Object Storage configuration interface for a bucket named "yslbckt12". The "Configuration" tab is active. Under the "Lifecycle policy" section, the "Static website hosting" tab is selected. The "Bucket details" section provides the following information:

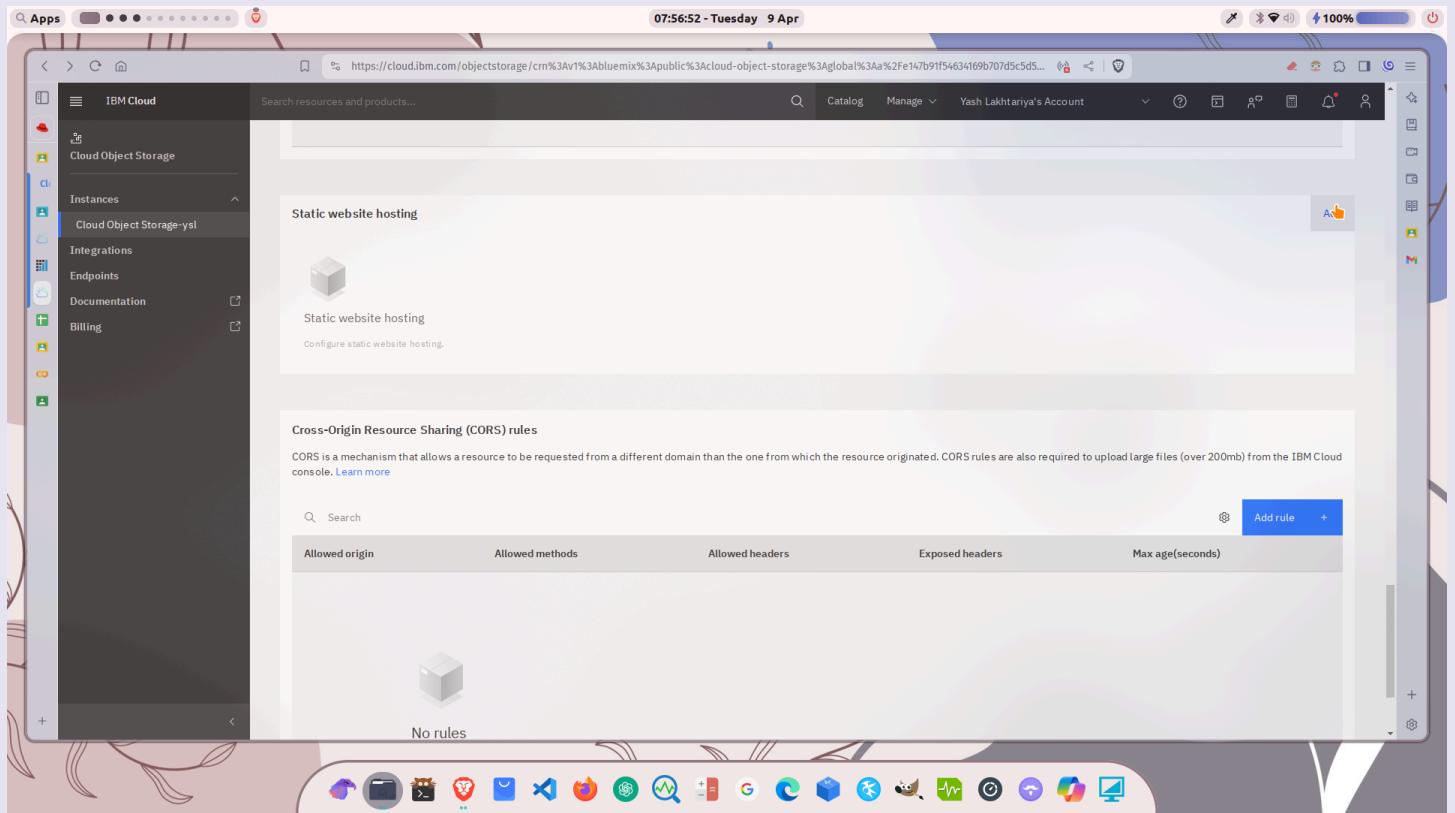
Bucket name	yslbckt12	Total bytes	1.2 MB
Service instance	cloud-object-storage	Resiliency	Cross Region
Total objects	1	Location	United States Geo (us-geo)
Storage class	Smart Tier	Date created	2024-04-03 8:43 AM
Cloud Functions trigger	Disabled Learn more		

The "Endpoints" section contains the Cloud Resource Name (CRN):

```
crn:v1:bluemix:public:cloud-object-storage:global:a:e147b91f64634169b707d5c5d51b9c77:c2b0aef9-875f-4e7c-a336-4397d8190135:bucket:yslbckt12
```

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

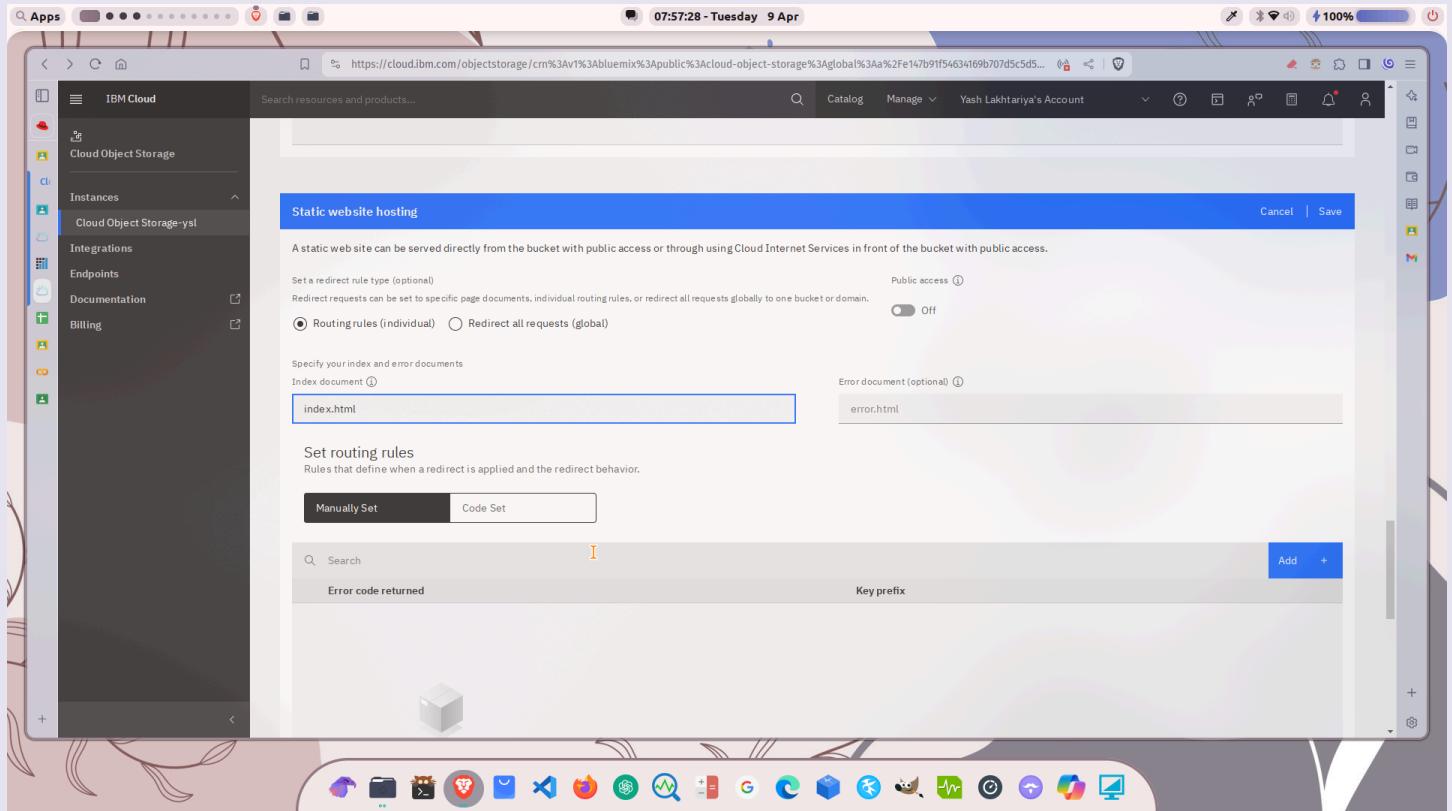
2. Select Add option in the tab



The screenshot shows the IBM Cloud Object Storage interface. On the left, there's a sidebar with options like 'Cloud Object Storage', 'Instances', 'Integrations', 'Endpoints', 'Documentation', and 'Billing'. The main area displays a 'Static website hosting' section with a small icon and a link to 'Configure static website hosting'. Below this is a 'Cross-Origin Resource Sharing (CORS) rules' section. A blue rectangular box highlights the 'Add rule' button, which is located at the top right of a table header. The table has columns for 'Allowed origin', 'Allowed methods', 'Allowed headers', 'Exposed headers', and 'Max age(seconds)'. The message 'No rules' is displayed below the table. At the bottom of the screen, there's a decorative footer bar with various icons.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

3. Save after entering the name html file



Name - Yash Lakhtariya

Enrollment number - 21162101012

Branch - CBA Batch - 61

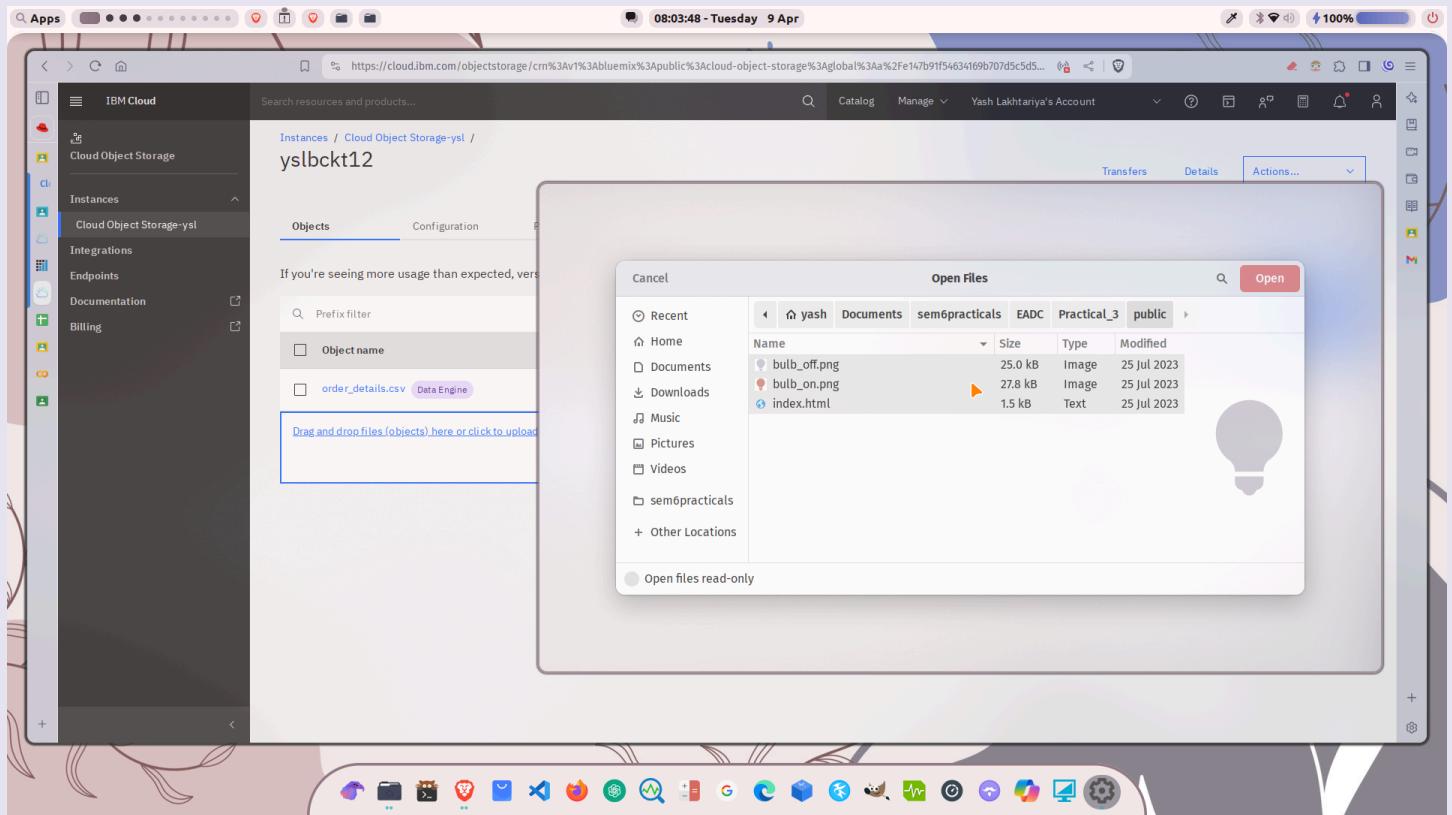
EADC Practical 12

- As seen the configuration is saved and URL is created for static website, but it won't work until the file is uploaded to object storage

The screenshot shows the IBM Cloud Object Storage interface. On the left, a sidebar lists 'Cloud Object Storage' under 'Instances', with 'Cloud Object Storage-ysl' selected. The main panel displays the 'Static website hosting' configuration for this bucket. It includes sections for 'Public access' (set to 'Off'), 'Index document' (set to 'index.html'), and 'Bucket website endpoint' (set to 'yslbckt12.s3-website.us.cloud-object-storage.appdomain.cloud'). Below these, there's a 'Set routing rules' section with a 'Manually Set' tab selected. A toolbar at the bottom contains various application icons.

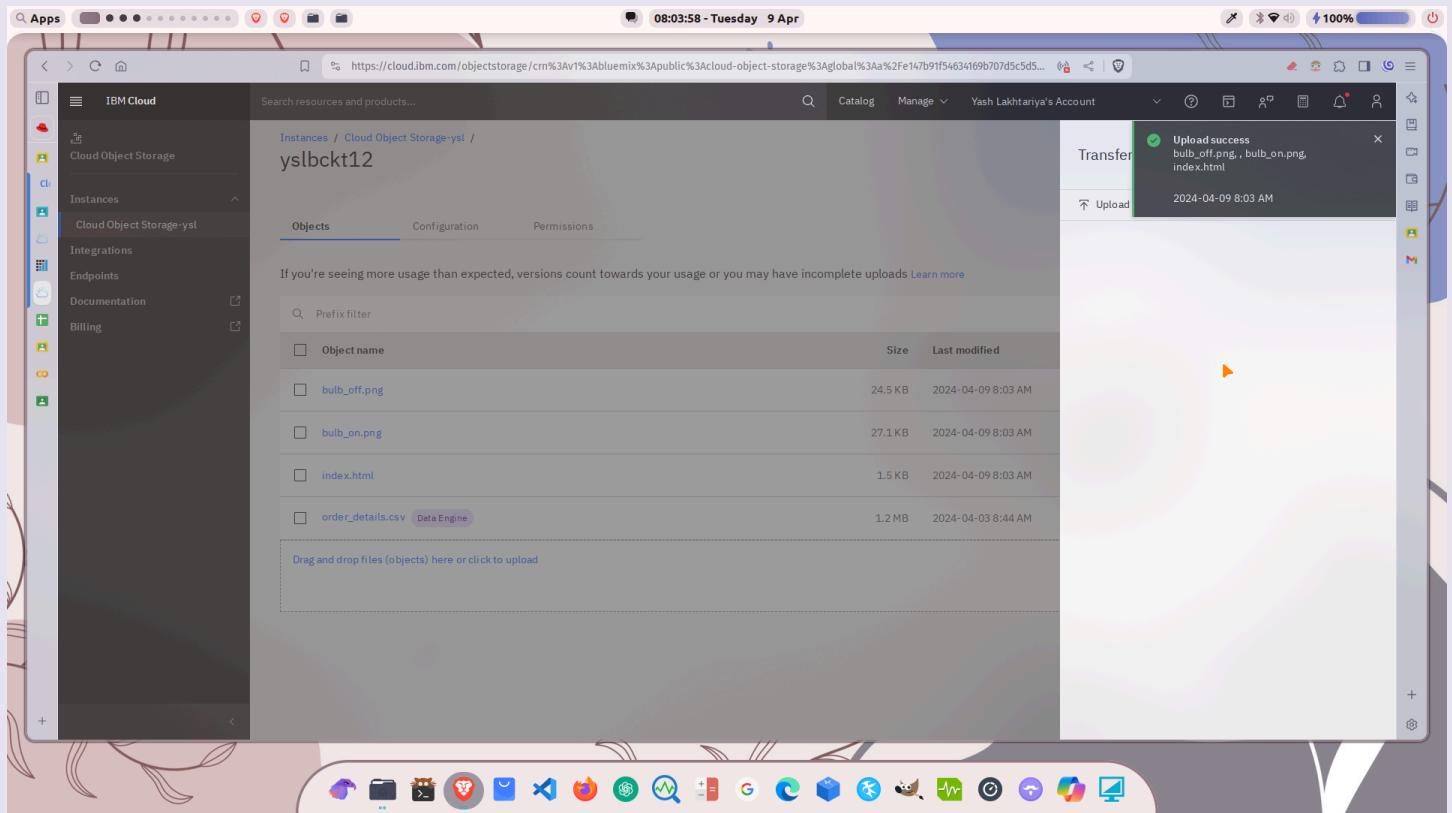
Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

5. Upload the required files in bucket



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

6. Wait till files are uploaded

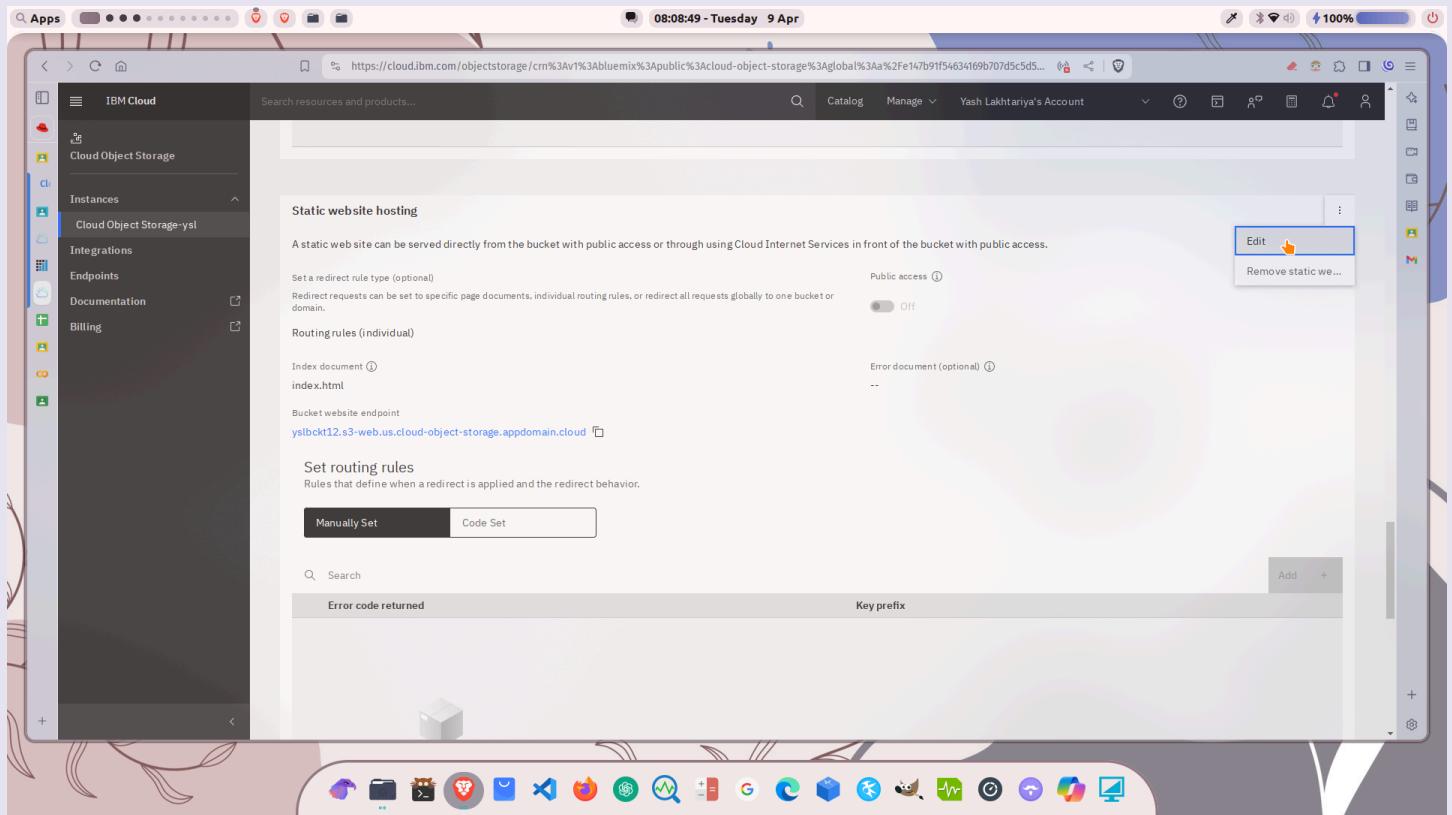


The screenshot shows the IBM Cloud Object Storage interface. The left sidebar lists 'Cloud Object Storage' under 'Instances'. The main area shows a list of objects in the 'Cloud Object Storage-ysl' instance, specifically in the 'yslbckt12' folder. The table displays four files: 'bulb_off.png', 'bulb_on.png', 'index.html', and 'order_details.csv'. A tooltip on the right indicates an 'Upload success' message for 'bulb_off.png, bulb_on.png, index.html' at '2024-04-09 8:03 AM'.

Object name	Size	Last modified
bulb_off.png	24.5 KB	2024-04-09 8:03 AM
bulb_on.png	27.1 KB	2024-04-09 8:03 AM
index.html	1.5 KB	2024-04-09 8:03 AM
order_details.csv	1.2 MB	2024-04-09 8:03 AM

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

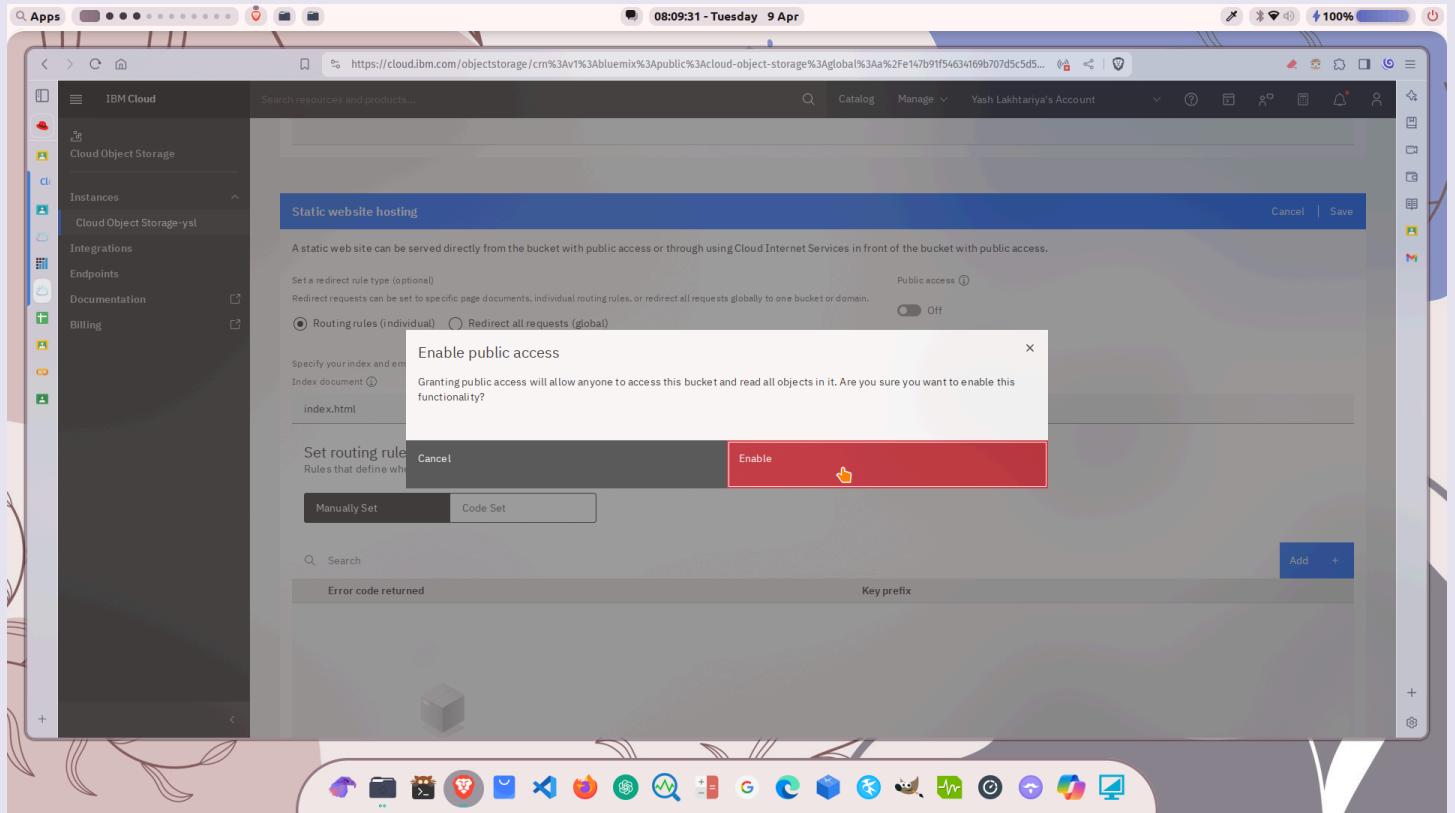
7. Now, to make website visible, edit the configuration



The screenshot shows the IBM Cloud Object Storage interface for managing a static website. On the left, a sidebar lists 'Cloud Object Storage', 'Instances' (selected), 'Cloud Object Storage-yasl', 'Integrations', 'Endpoints', 'Documentation', and 'Billing'. The main panel displays the 'Static website hosting' configuration for the 'Cloud Object Storage-yasl' instance. It includes sections for 'Set a redirect rule type (optional)', 'Public access' (set to 'Off'), 'Index document (optional)' (set to 'index.html'), 'Error document (optional) (optional)', 'Bucket website endpoint' (set to 'yslbckt12.s3-web.us.cloud-object-storage.appdomain.cloud'), and 'Set routing rules' (with tabs for 'Manually Set' and 'Code Set'). A search bar at the bottom allows for filtering 'Error code returned' and 'Key prefix'. A toolbar at the bottom contains various icons for file operations like upload, download, and search.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 12

8. Enable the public access and save the configuration



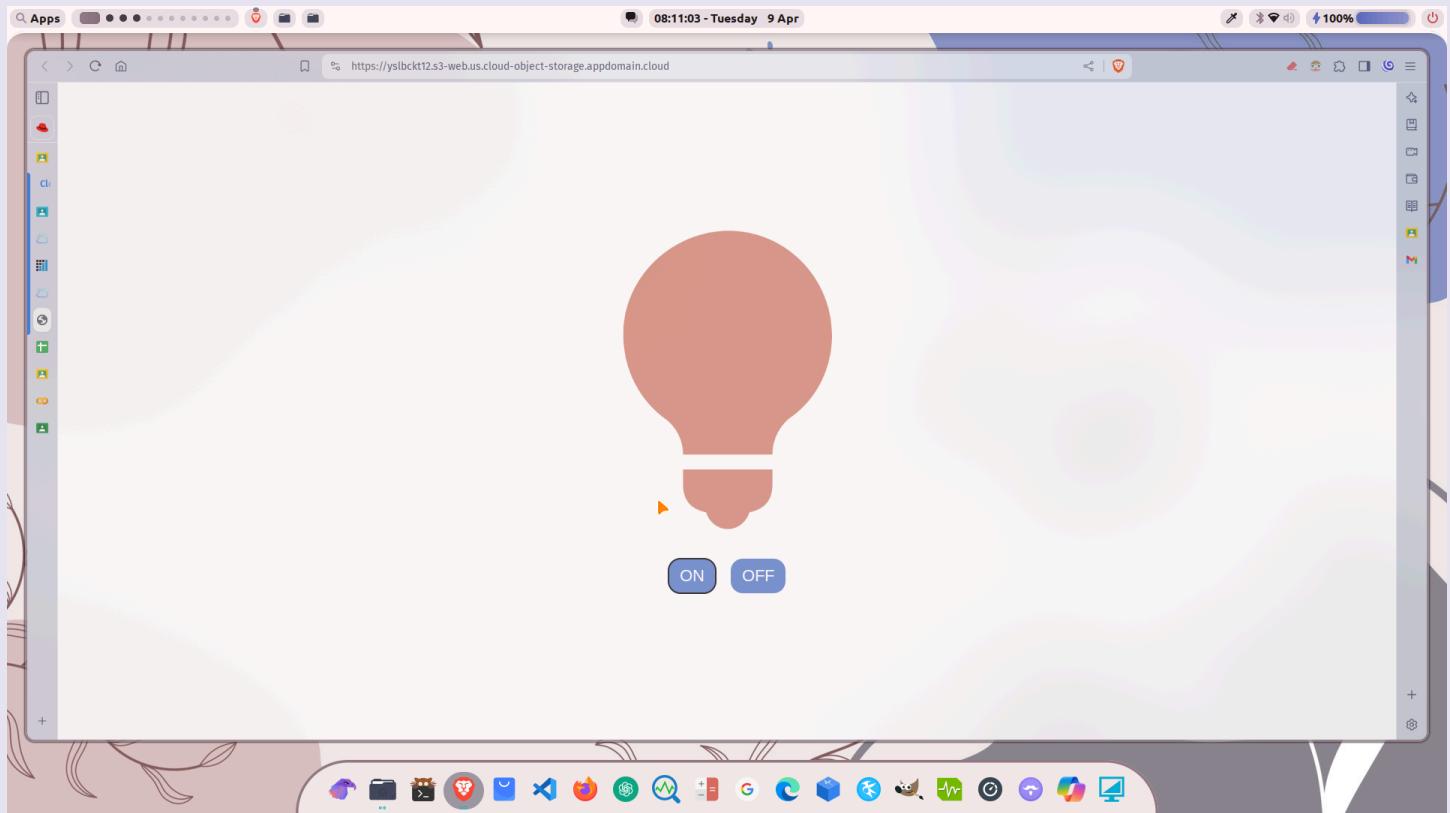
Name - Yash Lakhtariya

Enrollment number - 21162101012

Branch - CBA Batch - 61

EADC Practical 12

9. Now, visit the URL and the website should work as expected



URL : <https://yslbckt12.s3-web.us.cloud-object-storage.appdomain.cloud/>