

IBM CLOUD PAK FOR DATA 4.5

DATA VIRTUALIZATION HANDS ON LAB

Sandeep Ved
CUSTOMER SUCCESS MANAGER ARCHITECT, IBM SINGAPORE

Table of Contents

Data Virtualization Lab	3
Access Your Cluster.....	3
Service Provisioning	3
Data Virtualization (DV).....	4
DB2 (db2oltp).....	5
Jupyter Notebook	6
PostgreSQL	7
Manage Users Access	7
Db2 data loading	10
Check DB2 Connection Details.....	15
Setting up Data Virtualization.....	17
Task 1: Create a project.....	17
Task 2: Add a connection to your Db2 data source	19
Task 3: Add a connection to a PostgreSQL data source	22
Task 4: Add tables to your virtualized data	24
Task 5: Join and Publish virtualized data to project	27
Next steps.....	34
Analyze the data in a notebook	34
Visualize the data with a dashboard	36
Appendix.....	40
DB2 Instance Provisioning.....	40

Data Virtualization Lab

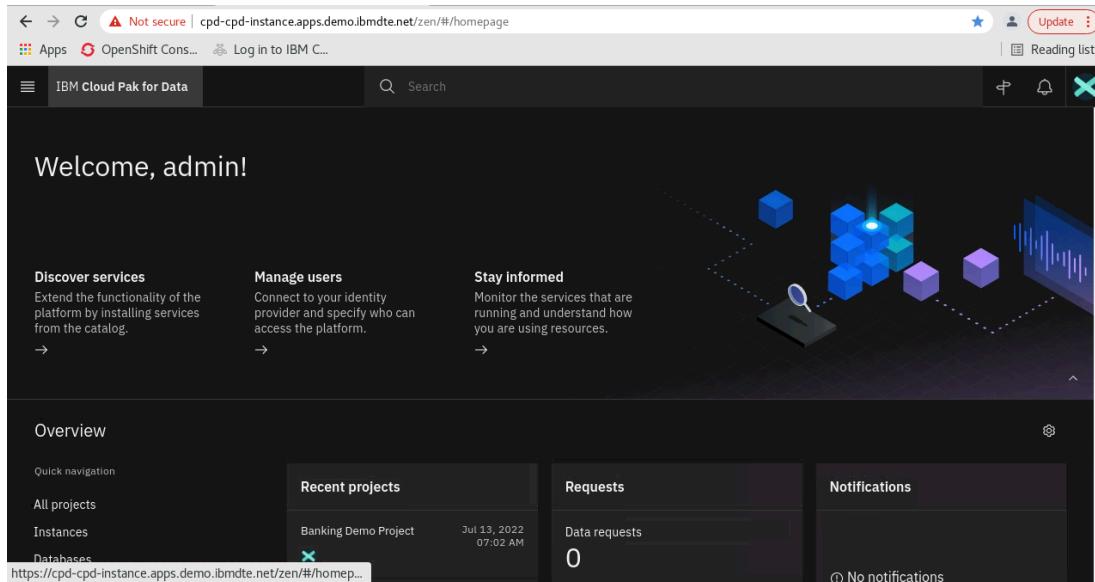
Access Your Cluster

- 1) Find out the OpenShift cluster Details as per the Given Cluster URL and credential provided.
- 2) Open the **Chrome/Firefox** Browser in **Incognito or Private mode** to Open the Cloud pak for Data Web Console.
- 3) Login to Web Console using the ID and credentials provided.

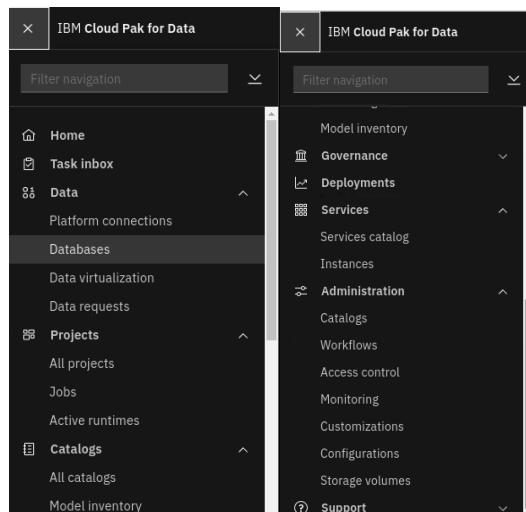
Service Provisioning

This lab requires the Data Virtualisation service as well as a Db2 service to prepare the data source.

- 1) Navigate back to CP4D Web Console Home Page.



- 2) Review all the Menu Items. Click on **Services -> Instances** to check the status of all instances.



- 3) Review the currently provided instances and its status. The status for all currently procured instances should be **green**.

Name	Type	Created by	vCPU requests	Memory requests (GiB)	Users	Status	Created on
openscale-defaultinstance IBM Watson OpenScale	aios	admin	0.00	0.00 Gi	1	Green	6 Jan 2023
data-management-console Data Management Console	dmc	admin	4.20	10.62 Gi	1	Green	5 Jan 2023

- 4) For the Lab, you need to create instances of Data Virtualization Instance (dv) and DB2 (db2oltp) services. The Data Management Console (dmc) service should already be enabled.
- 5) Click **New Instance +** Button to see what all services are already enabled and to create instance of dv and db2oltp. Review the status for all services. You will see that many services like Watson Studio or Cognos Dashboard are already enabled.

Data Virtualization (DV)

- 6) Select Data Virtualization Services.

Watson Knowledge Catalog	
IBM	Organize and govern data. Automatically discover, classify, profile, and protect your data so data scientists can find trusted data fast.
Enabled ✓	

Data sources 3 items

Data Virtualization IBM Query many data sources as one.	Db2 IBM Relational database that delivers advanced data management and analytics capabilities for transactional and warehousing workloads.	Db2 Data Management Console IBM Administer, monitor, manage, and optimize your Db2 databases from a single web-based console.
--	---	--

- 7) Click on the new Instance button and proceed the instructions.

The screenshot shows the IBM Cloud Pak for Data interface. At the top, there's a navigation bar with 'IBM Cloud Pak for Data' and a search bar. Below the navigation bar, the URL 'Instances / Select a service / Data Virtualization /' is visible. The main content area has a title 'Data Virtualization'. On the left, there's a sidebar with 'About' selected, showing details like Type: Application, Version: 1.8.3, Provider: IBM, Category: Data sources, and Related links. The main panel shows a 'Summary' section with the sub-section 'Virtualize'. It includes a table titled 'Available tables: 2521' with two rows:

Table	Schema	Database	Hostname: Port	Columns
CREDIT	ICPDCONTEST	tpods	9.30.167.128: 50000	3
TEMPLATE CONFIG	IBMCORSOLE	birtsol	9.30.160.36: 32051	5

DB2 (db2oltp)

8) Select DB2 Services.

The screenshot shows the 'Select a service' page. At the top, there's a navigation bar with 'IBM Cloud Pak for Data' and a search bar. Below the navigation bar, the URL 'Instances / Select a service /' is visible. The main content area has a title 'Select a service' with the subtitle 'Select the service for which you want to create a new instance.' Below this, there are several service cards. One card for 'Watson Knowledge Catalog' is shown in detail, while others for 'Data Virtualization', 'Db2', and 'Db2 Data Management Console' are partially visible.

Watson Knowledge Catalog
IBM
Organize and govern data. Automatically discover, classify, profile, and protect your data so data scientists can find trusted data fast.
Enabled ✓

Data sources 3 items

Data Virtualization IBM Query many data sources as one.	Db2 IBM Relational database that delivers advanced data management and analytics capabilities for transactional and warehousing workloads.	Db2 Data Management Console IBM Administer, monitor, manage, and optimize your Db2 databases from a single web-based console.
--	---	--

9) Click on the new Instance button and proceed the instructions. Refer to **appendix** to see the complete procedure for reference.

The screenshot shows the IBM Cloud Pak for Data interface. At the top, there's a navigation bar with 'IBM Cloud Pak for Data' and a search bar. Below it, a breadcrumb trail shows 'Instances / Select a service / Db2 /'. On the left, a sidebar provides details about the Db2 service, including its type (Database), version (11.5.7.0-cn7-x86_64), provider (IBM), category (Data sources), and related links (Docs). The main content area is titled 'Summary' and describes Db2 as a relational database for transactional and warehousing workloads. It includes a screenshot of the Db2 Advanced Edition VPC Option-5 interface, showing a table of SQL statements and their execution metrics.

Jupyter Notebook

10) Navigate to Menu -> Services -> Service Catalogue.

11) You can click on Jupyter Notebook and deploy the notebook.

The screenshot shows the 'Services catalog' page. At the top, there's a search bar and a sidebar with categories like 'Category', 'Entitlement', 'Source', and 'Status'. Below the sidebar, there are sections for 'Explore more services' (Partners catalog), 'Industry accelerators' (Jump-start your analysis of common business problems), and 'Services catalog' (Find services). The 'Developer tools' section is highlighted with a blue border and contains three items: 'Jupyter Notebooks with R 3.6' (optional development environment for Watson Studio), 'Lightbend Platform' (Reactive Microservices, real-time streaming, and Machine Learning), and 'RStudio Server with R3.6' (optional development environment for Watson Studio). The 'Industry solutions' section contains two items: 'Financial Services Workbench' (design, implement, and deliver microservices) and 'Prolifics Customer Prospecting Accelerator' (unlock reliable sales leads).

12) Click on Deploy button and continue the instructions

The screenshot shows the IBM Cloud Pak for Data Practicum interface. At the top, there's a navigation bar with 'IBM Cloud Pak for Data' and a search bar. Below the navigation bar, a breadcrumb trail shows 'Services catalog /'. On the right side of the header, there are icons for 'Deploy' and a user profile.

The main content area has a title 'Jupyter Notebooks with R 3.6'. Below the title, there are tabs for 'IBM' and 'Docs', with 'Docs' being the active tab. A sidebar on the left is titled 'About' and contains sections for 'Type Application', 'Installation dependencies Watson Studio', 'Provider IBM', 'Category Developer tools', and 'Related links Docs'.

The central area displays a Jupyter Notebook interface with a summary section and a code editor. The summary section describes it as an optional development environment for Watson Studio that enables you to create Jupyter Notebooks that use R 3.6 libraries. The code editor shows a notebook titled 'Use R to forecast time series data'.

PostgreSQL

- 1) An external PostgreSQL DB database has been already prepared for the lab with the following connection details.
 - Database Name: vlyawtap
 - Host Name: tiny.db.elephantsql.com
 - Port: 5432
 - User Name: vlyawtap
 - Password: fXt4TyCB_W0d0LCaCaPF7MbLKWlpti60

Manage Users Access

- 1) Navigate to CP4D Console Home Page. Go to Menu -> Services -> Instances.

The screenshot shows the CP4D Console Home Page. At the top, there's a navigation bar with 'Not Secure | https://cpd-cpd-instance.apps.itzocp-664003qtc-mn9o.selfservice.aws.techzone.ibm.com/zen/#/homepage'. Below the navigation bar, a 'Welcome, admin!' message is displayed.

The main content area features three columns: 'Discover services', 'Manage users', and 'Stay informed'. The 'Stay informed' column includes a 'Cloud Data Platform' graphic showing a cluster of cubes connected by lines.

On the left, there's a sidebar with 'Overview' and 'Quick navigation' sections, including links for 'All projects', 'Instances', 'Databases', 'Data virtualization', 'Support', 'Documentation', 'Community', and 'Diagnostics'.

The main dashboard area has sections for 'Recent projects', 'Requests', and 'Notifications'. The 'Recent projects' section shows a project named 'zz01jj856' from Sep 07, 2022, at 10:59 PM. The 'Requests' section shows 'Data requests 0'. The 'Notifications' section shows two notifications: 'Online deployment ready' for 'New model' in space 'Credit Risk' and 'Online deployment ready' for 'Credit Risk Model - Random Forest Classifier - Deployment ii'.

2) For the *Db2 instance*, and select **Manage Access**.

Name	Type	Created by	vCPU requests	Memory requests (GiB)	Users	Status	Created on
data-virtualization	dv	admin	11.50	38.50 Gi	1	Green	13 Sept 2022
data-management-console	dmc	admin	5.20	21.40 Gi	1	Green	8 Sept 2022
Db2-1 Service Instance for db2oltp-1662568821240436	db2oltp	admin	0.60	4.25 Gi	1	Green	8 Sept 2022
openscale-defaultinstance	aios	admin	-	-	1	Green	7 Sept 2022

3) Review the users with Access to DB2.

Name	Username	Service role
admin	admin	Admin

4) You may add new team members users by clicking on **Add User+** button for them to have access to manage DB2 Services.

Grant access to users and user groups
Specify the users who can access Db2 and the role of each user.

Users and user groups	Filter by:
admin	All
Sandeep Ved	Admin

Add

Name	Username	Service role
admin	admin	Admin
Sandeep Ved	sved	Admin

5) Similarly assign access to new team members to the data virtualization instance as well.

Name	Type	Created by	vCPU requests	Memory requests (GiB)	Users	Status	Created on
openscale-1 IBM Watson OpenScale	aios	admin	0.00	0.00 Gi	1	Green	20 Sept 2022
data-virtualization	dv	admin	11.50	38.50 Gi	3	Green	13 Sept 2022
data-management-console IBM Db2 Data Management Console	dmc	admin	5.20	21.40 Gi	3	Green	8 Sept 2022
Db2-1 Service instance for db2oltp-16625688212...	db2oltp	admin	0.60	4.25 Gi	3	Green	8 Sept 2022

Name	Username	Service role
admin	admin	Admin
Satit Pongbundit	satit.pongbundit	Engineer
Sandeep Ved	sved	Admin

6) Similarly assign access to new users to the data management console instance as well.

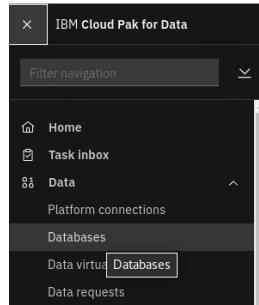
Name	Type	Created by	vCPU requests	Memory requests (GiB)	Users	Status	Created on
openscale-1 IBM Watson OpenScale	aios	admin	0.00	0.00 Gi	1	Green	20 Sept 2022
data-virtualization	dv	admin	11.50	38.50 Gi	3	Green	13 Sept 2022
data-management-console IBM Db2 Data Management Console	dmc	admin	5.20	21.40 Gi	3	Green	8 Sept 2022
Db2-1 Service instance for db2oltp-16625688212...	db2oltp	admin	0.60	4.25 Gi	3	Green	8 Sept 2022

Name	Username	Service role
admin	admin	Admin
Sandeep Ved	sved	Admin
Satit Pongbundit	satit.pongbundit	User

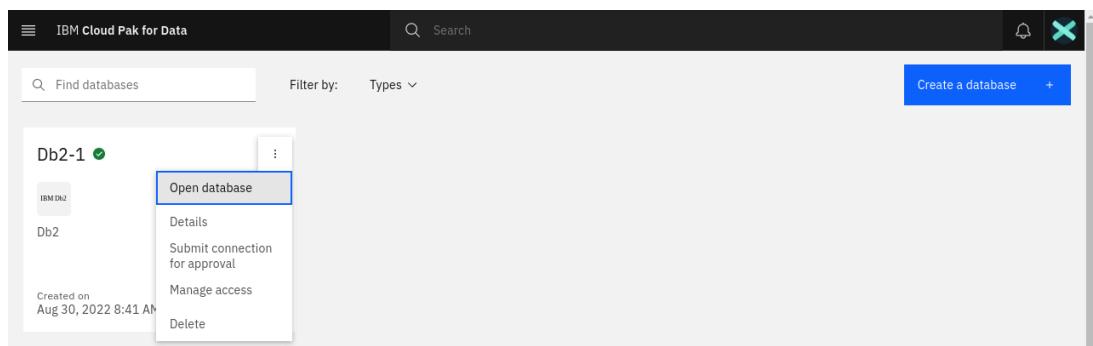
Db2 data loading

The data sets you will use in this tutorial are available in the Practicum Github. You need to load those data sets into Db2.

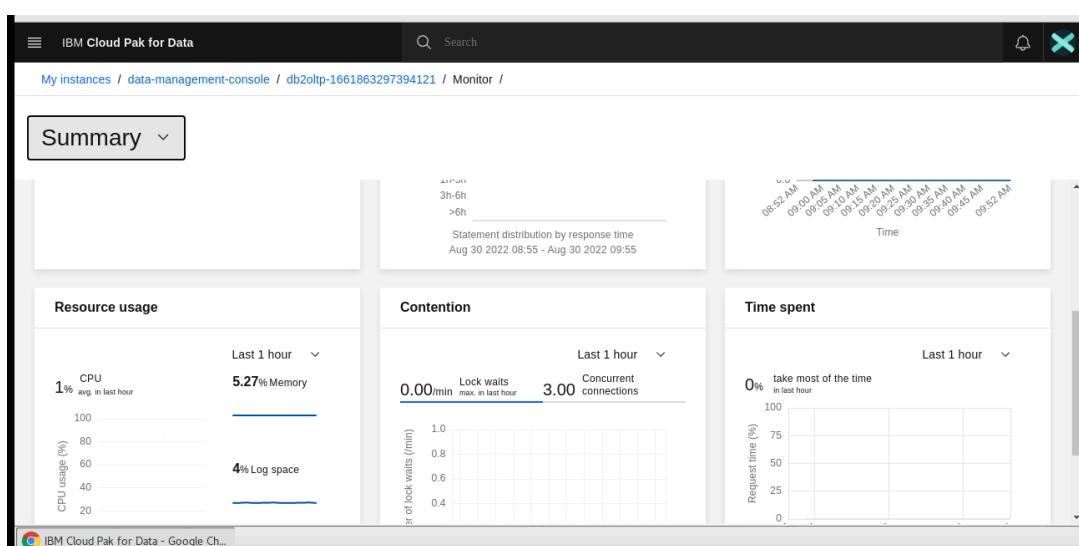
1. Access these two data sets in the GitHub, and download them to your local machine:
 - o **Customers data set -> customers.csv** (To upload data in DB2 Database)
 - o **Sales data set -> sales.csv** (For Reference only, as its already loaded in PostgreSQL DB)
2. Navigate to **Data -> Databases.**



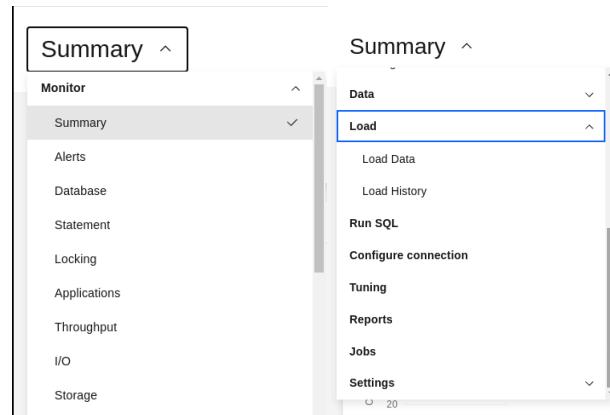
3. Click on **Open Database**



4. You will see the DB2 Database Summary. Click on the Summary Drop Down.



5. Under the Summary Drop down, click on **Load -> Load Data.**



6. The DB2 Load Data Page will appear. Select or Drag the **customers.csv** file from your local machine into the *File Selection* panel in the *Db2* console.

The screenshot shows the 'Load Data' page in the Db2 console. The 'Source' tab is selected, showing a 'File selection' panel with a 'Drag a file here or browse files' area and a 'File selection' button. Below it is an 'Open File' dialog box showing the file 'Customers.csv' selected in the 'Downloads' folder. The 'Target' and 'Define' tabs are also visible at the top.

7. Click **Next**.

The screenshot shows the 'Load Data' page again. The 'Selected file' panel on the right now lists 'Customers.csv'. At the bottom right of the page, there is a blue rectangular button labeled 'New Schema+'.

8. Click **New Schema+**.

Source Target Define Finalize

You are loading the file **Customers.csv**

Select a load target

Schema

Find schemas

TUTORIAL
aiopenscale00

Back Next

9. Enter a suitable name, eg. IBMADMIN. Click **Create**.

Source Target Define Finalize

You are loading the file **Customers.csv**

Select a load target

Schema

Find schemas

TUTORIAL
aiopenscale00

Create a new schema

IBMADMIN

Create

Back Next

10. Click **New table +**.

Load Data

Source Target Define Finalize

You are loading the file **Customers.csv**

Select a load target

Schema Table

IBMANDMIN No entries found.

Back Next

11.Type **customers** for the table name, and click **Create**.

Load Data

Source Target Define Finalize

You are loading the file **Customers.csv**

Select a load target

Schema Table

IBMANDMIN No entries found.

Create a new table

Customers

Create

Back Next

12.Click **Next** to continue.

Source Target Define Finalize

You are loading the file **Customers.csv** into **IBMANDMIN.CUSTOMERS**

Select a load target

Schema **Table**

New schema + New table +

Find schemas Find tables in IBMANDMIN

IBMANDMIN CUSTOMERS

IBMANDMIN TUTORIAL aiopenscale00

Back Next

13. Review the data to be loaded, and click **Next**.

Source Target Define Finalize

You are loading the file **Customers.csv** into **IBMANDMIN.CUSTOMERS**

Code page (character encoding): 1208 (UTF-8) Separator: ,

CUST_ID	CUSTNAME	CITY	STATE	COUNTRY_CODE	POSTAL_CODE	EMAIL_ADDRESS	PHONE_NUMBER	YTD_SALES
1	Michael Golden	Abbeida Alpina	TO	IT	10060	Michael.E.golden@spambob.com	724-454-8453	90.3
2	Renee Mullins	Columbus	OH	IT	45101	Rene.K.mullins@dodgeit.com	229-990-2162	0
3	Allen Schmidt	Abeta	PG	IT	6040	Allen.M.schmidt@spambob.com	289-202-8653	304
4	Robert May	Houston	TX	US	77601	Robert.C.may@spambob.com	630-492-6553	304
5	Joe Cruz	Harrisonburg	VA	US	24210	Joe.D.cruz@spambob.com	449-788-1089	180.3
6	Rebecca White	Acate	RG	IT	97011	Rebecca.white@spambob.com	826-268-4303	52
7	Gary Neal	Chicago	IL	US	61701	Gary.N.neal@dodgeit.com	267-929-9454	673
8	Steve Huynh	Detroit	MI	US	48614	Steve.A.huynh@dodgeit.com	644-191-7549	904.86
9	Anthony Johnson	Achstetten	DE	DE	88480	Anthony.H.johnson@pookmail.com	907-111-5490	354
10	Alberto Fabian	Acquabona	BL	IT	32043	Alberto.T.fabian@pookmail.com	113-785-7672	352

Back Next

14. Review the summary, and click **Begin Load**.

Source Target Define Finalize

You are loading the file **Customers.csv** into **IBMANDMIN.CUSTOMERS**

Review settings

Summary

Code page: 1208 (Default)
Separator: , (Default)
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

Back Begin Load

15. Verify that the load is successful. Click **View Table**.

The screenshot shows the 'Load Data' interface for IBM Cloud Pak for Data. It displays a summary of the load operation: 100 rows read, 100 rows loaded, and 0 rows rejected. The status bar indicates the start time as 08/30/2022 10:14:30 AM and the end time as 08/30/2022 10:14:43 AM. A large blue circular progress indicator is present. To the right, a message states 'The data load job succeeded.' and 'You can now work with your data.' Below this, an 'Errors' section shows 'No errors'.

16.Table details will appear.

The screenshot shows the 'IBADMIN.CUSTOMERS' table details page. It displays a grid of 12 rows of customer data with columns including CUST_ID, CUSTNAME, CITY, STATE, COUNTRY_CODE, POSTAL_CODE, EMAIL_ADDRESS, PHONE_NUMBER, YTD_SALES, SALESREP_ID, NATIONALITY, NATIONAL_ID, CREDITCARD_TYPE, CREDITCARD_NUMBER, and CREDITCARD_SMALLNAME. The data includes various names like Michael Golden, Renee Mullins, and Robert May, along with their respective details such as city (Abetia, Alpha, Columbus), state (OH, US, TX), and country code (IT, US, TX).

17.Repeat steps 5-10 to load the **sales.csv** data set into the same schema with the table name **sales**.
<not needed as same table is uploaded and available from PostgreSQL DB>

Check DB2 Connection Details

18.Navigate back to Data -> Databases from the main navigation menu.

The screenshot shows the 'Databases' interface. It lists a single database entry: 'Db2-1' with a green status icon. Below it, there are two sub-folders: 'IBM DB2' and 'Db2'. A note at the bottom states 'Created on Aug 30, 2022 8:41 AM'.

19.Click on Details .

The screenshot shows the 'Details' menu for the 'Db2-1' database. The 'Open database' option is highlighted with a blue border. Other options include 'Submit connection for approval', 'Manage access', and 'Delete'.

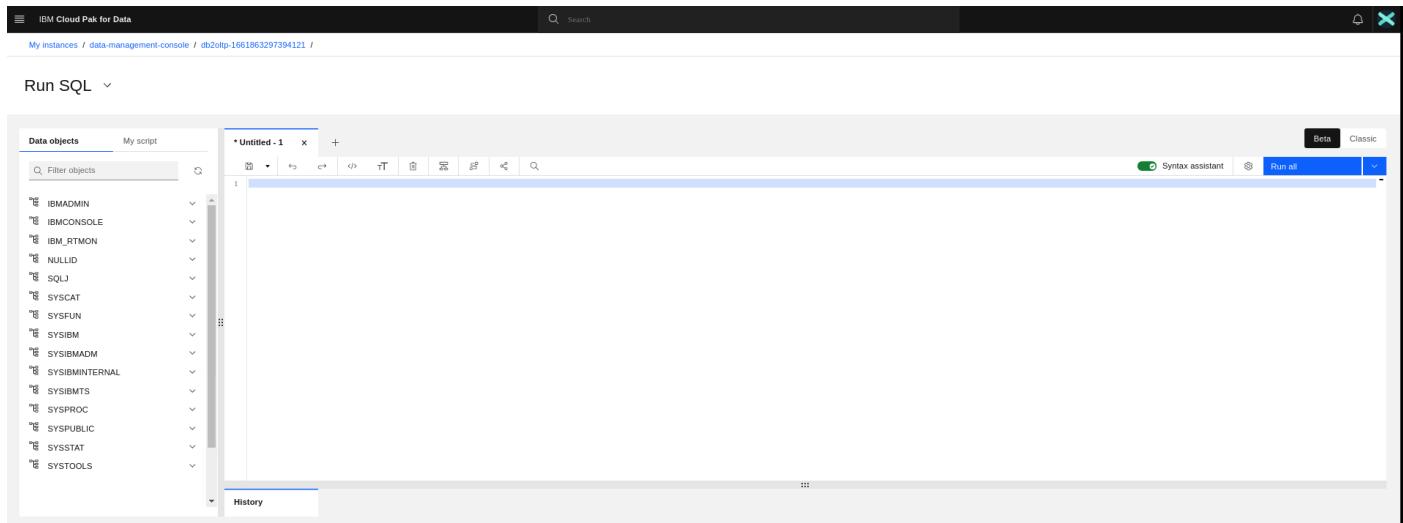
20. Note down the connection details for this DB instance like hostname, Database Name, Port (non SSL & SSL). You can use the admin user id credential to connect to the database later.

The screenshot shows the 'Details' page for the 'Db2-1' database. It includes sections for 'About this database' (with fields like Database name: BLUDB01, Database type: db2oltp, etc.) and 'Access information' (with JDBC Connection URLs for non-SSL and SSL). A 'Nodes' table lists a single node: ip-10-0-132-180.ap-southeast-1.compute.internal. A 'Download SSL Certificate' button is also present.

21. Again open the DB2 Database and Open the Summary Menu and select Run SQL.

The screenshot shows the 'Authorization' interface under the 'Data' section of the summary menu. It displays a list of roles and users, with buttons for 'Add role', 'Remove role', 'Grant multiples', and 'Revoke multiples'. A search bar and a 'Find name' input field are also visible.

22. You can use the SQL interface to update the DB2 database as usual.



23. For example you can run below commands to grant full access to the database to newly created users.

Script	Date	Status	Runtime
grant dbadm on database to user sved	Sep 16, 2022 3:18:05 PM	✓ 2	0.012 s
grant secadm on database to user sved	Sep 16, 2022 3:18:05 PM	✓	0.006 s
	Sep 16, 2022 3:18:05 PM	✓	0.006 s

Setting up Data Virtualization

You can use **Data Virtualization** to create a virtual table to segment or combine data from one or more tables. **Data Virtualization** connects multiple data sources into a single self-balancing collection of data sources or databases. Refer this link for more details -> <https://www.ibm.com/docs/en/cloud-paks/cp-data/4.5.x?topic=data-virtualizing>

Task 1: Create a project

You need a project to store the virtualized data.

1. Login to CPD Platform. Click on the Hamburger Icon on the top left.
2. Select **View all Projects**. Alternatively All Projects can be accessed from Home Page under Overview Section.
3. If you have an existing project and want to reuse it, open it. If you don't have an existing project or want to use fresh new project, click **New project** on your **Projects** page.

4. Select **Create an empty project**.

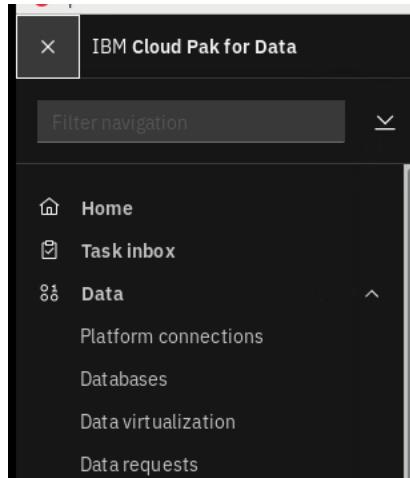
5. On the **Create a project** screen, add a name and optional description for the project. Click **Create**.

6. The Project will be created successfully.

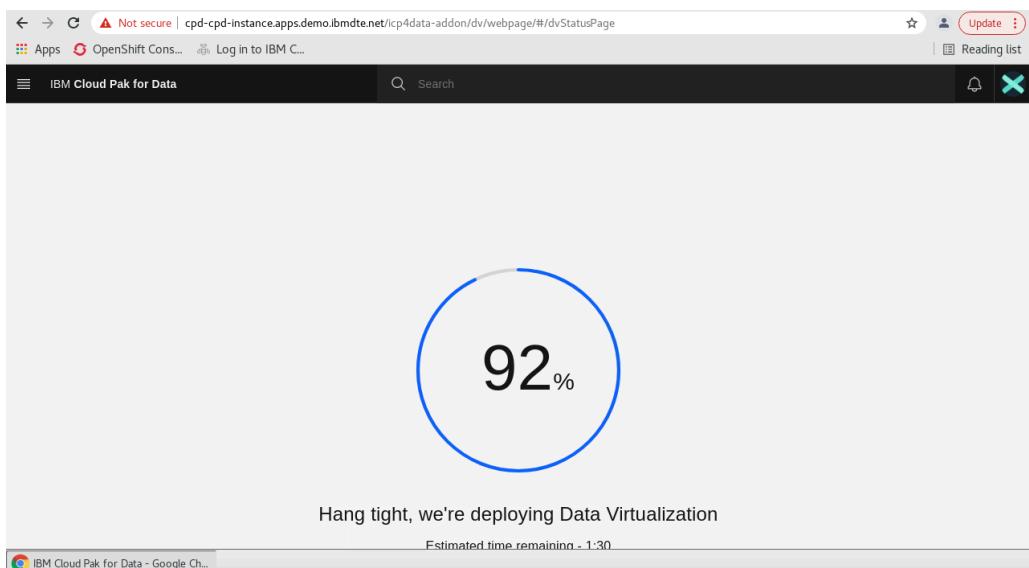
Task 2: Add a connection to your Db2 data source

Before you can virtualize the data, you need to create a connection to the data source.

1. From the main navigation menu, select **Data > Data virtualization**.



2. If opening for the first time, the Data Virtualization component may start to get initialized. Just wait until its fully deployed and Next screen appears.



3. After initial initialization, the Data Virtualization Data Sources Page appears.

The screenshot shows the 'Data sources' section of the IBM Cloud Pak for Data interface. At the top right, there is a blue 'Add connection' button. Below it, there are two tabs: 'Table view' (selected) and 'Constellation view'. A search bar is located above the table. The table has columns: Name, Endpoint, Type, Status, and Listed tables. There is one row in the table with a question mark icon and a person icon. The status is 'Unknown'.

4. Click Add connection > New connection.

The screenshot shows the 'New connection' dialog box. The 'Add connection' tab is selected. It has three options: 'New connection' (selected), 'Existing platform connection', and 'Remote data source'. Below the tabs, there is a message: 'You don't have any data sources yet. Add a data source by creating a new connection or select from an existing platform connection or by searching for a data source via a remote connector.'

5. Select IBM Db2. Click Select.

The screenshot shows the 'New connection' dialog box with 'IBM Db2' selected in the provider dropdown. The 'Selected connection type' section shows 'IBM Db2' is selected. The 'Details' section lists 'IBM Db2 database'. The 'Compatible services' section includes Catalogs, DataStage, Data Virtualization, Data discovery and Data quality, Metadata import (lineage), Metadata import (discovery), and Watson Studio. At the bottom, there are 'Cancel' and 'Select' buttons, with 'Select' being highlighted in blue.

6. Complete the connection details based on the credentials you copied for your db2 instance .

Create connection: Db2
Enter the connection information.

Connection overview

Name: CP4D DB2 Connection

Description: CP4D DB2 Connection for Customers Data

The details for this type of connection cannot be validated until you use the connection.
Ensure that the credentials and details are correct. If you encounter an issue, you can edit the connection later.

Connection details

Additional properties:

Back **Create**

7. Enter Database Name, Host Name and Non SSL Port as we noted earlier. Select the Credential Setting as **Personal**.

Create connection: Db2
Enter the connection information.

Connection overview

Connection details

Database*: Customer

Hostname or IP address*: worker1

Port*: 30812

Credentials

Credential setting: Personal Shared

Cloud Pak for Data authentication

Back **Create**

8. Enter **Personal** Credential of **admin** User. Click **Create**.

Create connection: Db2
Enter the connection information.

Credentials

Personal Shared

Cloud Pak for Data authentication

Username*: admin

Password*:

Username and Password Security Mechanism: Default

Username and Password Encryption Algorithm: Default

Certificates

Port is SSL-enabled

Back **Create**

9. Click **Skip**

You haven't set up any remote connectors yet.
You must set up a remote connector to assign data sources to the connector. You must have the required permissions to set up remote connectors.

[Set up remote connector](#)

10.The DB2 Data Source will be added as a Data Virtualization Source.

Name	Endpoint	Type	Status	Listed tables
CP4D DB2 Connection	worker1: 30812	Db2 Family	Active	318 / 318
Database	Port	Username	Remote connector	Description
Customer	30812	admin	-	CP4D DB2 Connection for Customers Data

Task 3: Add a connection to a PostgreSQL data source

For the purpose of Validating the Data Virtualization from Multiple Types of Data Sources, We have prepared a **PostgreSQL DB** as well with the *Sales.csv* data uploaded in table named *Sales_Rep*. Use below steps and given credentials to use the data from this PostgreSQL Data Source.

1. From the main navigation menu select **Data > Platform Connections**. The list of configured *Data sources* displays.
2. Click **New connection**.

The screenshot shows the 'Data sources' section of the IBM Cloud Pak for Data interface. At the top right, there are buttons for 'Set up remote connector', 'Add connection', 'New connection', 'Existing platform connection', and 'Remote data source'. A feedback survey overlay is visible. Below the header, a table lists one connection:

Name	Endpoint	Type	Status	Listed tables
CP4D DB2 Connection	worker1: 30812	Db2 Family	Active	318 / 318
Database	Port	Username	Remote connector	Description
Customer	30812	admin	-	CP4D DB2 Connection for Customers Data

3. Select PostgreSQL and click Select.

The screenshot shows the 'New connection' dialog. On the left, a sidebar lists providers: IBM, Third-party, and User-defined. The main area shows a search bar with 'postgres' and a list of connection types: 'All connection types', 'Amazon RDS for PostgreSQL', 'Databases for PostgreSQL', and 'PostgreSQL'. 'PostgreSQL' is selected and highlighted. On the right, details about the selected connection type are shown, including 'Selected connection type: PostgreSQL', 'Details: PostgreSQL database', and 'Compatible services: Catalogs, DataStage, Data Virtualization, Metadata Import, Metrics, Quality'. At the bottom are 'Cancel' and 'Select' buttons, with 'Select' being highlighted.

4. Enter the Connection Name and Description.

The screenshot shows the 'Create connection: PostgreSQL' dialog. On the left, a sidebar has tabs for 'Connection overview', 'Connection details', 'Credentials', and 'Certificates', with 'Connection overview' selected. The main area shows 'Connection overview' with fields for 'Name' (PostgreSQL on ElephantsQL) and 'Description' (Added by Sandeep Ved - Testing). At the bottom are 'Cancel' and 'Create' buttons, with 'Create' being highlighted.

5. Enter the Connection Details and Credentials as below. Click Next.

- Database Name: vlyawtap
- Host Name: tiny.db.elephantsql.com
- Port: 5432
- Credential: Personal
- User Name: vlyawtap
- Password: fXt4TyCB_W0d0LCaCaPF7MbLKWlpti60

Create connection: PostgreSQL

Enter the connection information.

Connection overview	Database* ⓘ vlywtap
Connection details	Hostname or IP address* ⓘ tiny.db.elephantsql.com
Credentials	Port* ⓘ 5432
Certificates	
Credentials	Credential setting ⓘ <input checked="" type="radio"/> Personal <input type="radio"/> Shared Username* ⓘ vlywtap Password* ⓘ <input type="button" value="Show password"/>
Certificates	<input type="checkbox"/> Port is SSL-enabled ⓘ

[Cancel](#) [Back](#) [Create](#)

6. Click Skip.

My instances / data-management-console / dv-1649188282576777 /

Add to a remote connector (optional)

You can add connections to remote connectors to enhance parallelism during processing and improve query performance.

Remote connector	Hostname	Description	Port	Username

You haven't set up any remote connectors yet.

[Skip](#) [Add to connector](#)

7. The new data source will be added.

IBM Cloud Pak for Data

My instances / data-management-console / dv-1857090051946032 / Virtualization /

Data sources

Table view Constellation view

Connections: 2 | Remote connectors: 0

Find by Name, Endpoint, Type

Name	Endpoint	Type	Status	Listed tables
PostgreSQL db ElephantsQL	tiny.db.elephantsql.com: 5432	PostgreSQL	Active	2 / 2
CP4D DB2 Connection	worker1: 30012	Db2 Family	Active	318 / 318

[Set up remote connector](#) [Add connection](#)

Task 4: Add tables to your virtualized data

With the connection defined, you can virtualize data from that data source.

- From the *Data sources* menu, select **Virtualization > Virtualized Data**.

The screenshot shows the 'Data sources' section of the 'Virtualization' tab. The left sidebar lists 'Virtualization', 'Data sources', 'Virtualize', 'Virtualized data', 'Cache management', 'Monitor', 'Data', 'Run SQL', and 'Administration'. The right panel shows two entries under 'Endpoint': 'tiny.db.elephantsql.com: 5432' and 'worker1: 30812'.

2. Click on Add Virtual Objects.

The screenshot shows the 'Virtualized data' section. It includes a search bar, filter options ('All types'), and a table header with columns 'Table', 'Schema name', 'Created on', and 'Statistics last collected on'. A message at the top states: 'You don't have any virtualized objects yet.' Below it, a note says: 'Add connections to your data sources, then you can create virtual tables and views. Learn more'. A blue 'Add virtual objects' button is visible at the bottom left.

3. You will see the available connections from where data can be virtualized.

The screenshot shows the 'Virtualize' section with the 'Tables' tab selected. The 'Connections' tab is open, showing a list with 'PostgreSQL db ElephantSQL...' and 'CP4D DB2 Connection (318)'. The 'PostgreSQL db ElephantSQL...' entry is selected. A message in the center says: 'You haven't made a selection yet. Select a connection to browse schemas and find tables that you want to virtualize.'

4. Select CP4D DB2 Connection -> IBMADMIN -> CUSTOMERS Table and click Add to cart.

The screenshot shows the 'Virtualize' section with the 'Tables' tab selected. The 'Connections' tab is open, showing 'CP4D DB2 Connection / IBMADMIN /'. The 'CUSTOMERS' table is selected and highlighted in blue. A message in the center says: 'You haven't made a selection yet. Select the tables that you want to virtualize and add them to the cart. Click a table to preview the table contents or click the checkbox to add it to the cart without previewing it.'

5. Click on Connections link again.

6. Select PostgreSQL on ElephantSQL -> sample -> sales_rep tables from the list, and click Add to cart.

7. Click View cart.

8. Select Assign to Virtualized data to add these two tables to your list of virtualized data. Click Virtualize.

Table	Schema	Source schema	Connections	Databases/File Path	Hostname: Port	Grouped tables
CUSTOMERS	ADMIN	IBADMIN	CP4D DB2 Connection	Customer	worker1: 30812	1
SALES_REP	ADMIN	sample	PostgreSQL on Elepha...	vfyawtap	tinydb.elephantsql.co...	1

9. Click Confirm.

Confirm virtualization

If you click Confirm and start virtualizing, you can't cancel the request or leave this screen until the virtualization is complete. If you click Cancel, your object will not be virtualized.

Do not show this message again

Cancel

Confirm

10. Click **Go to virtualized data**.

Table	Schema	Virtualization status	Publish status
CUSTOMERS	ADMIN	Success	Success
SALES_REP	ADMIN	Success	Success

Estimated virtualization time: 2 seconds

Virtualize more data Go to virtualized data

11. You can see the **virtualized data**.

Table	Schema name	Created on	Statistics last collected on
SALES_REP	ADMIN	Sep 2, 2022 11:21 AM	Not collected
CUSTOMERS	ADMIN	Sep 2, 2022 11:21 AM	Not collected

Task 5: Join and Publish virtualized data to project

Next, join two tables to create a virtualized asset and publish that to a catalog and project.

1. On the *Virtualized data* screen, select the **customers** and **sales_rep** tables from the list, and click **Join**.

Table	Schema name	Created on	Statistics last collected on
SALES_REP	ADMIN	Sep 2, 2022 11:21 AM	Not collected
CUSTOMERS	ADMIN	Sep 2, 2022 11:21 AM	Not collected

2. Connect the **SALESREP_ID** columns in the two tables. Click the link **Open in SQL Editor** to see the actual query used to join the tables and create the view.

Table 1: SALES_REP

<input checked="" type="checkbox"/> Column name	Data type
Salesrep_id	CLOB
First_name	CLOB
Last_name	CLOB
Nationality	CLOB
National_id	CLOB
Phone_number	CLOB
Age	CLOB
Sex	CLOB
Territory	CLOB

Table 2: CUSTOMERS

<input checked="" type="checkbox"/> Column name	Data type
State	VARCHAR
Country_code	VARCHAR
Postal_code	VARCHAR
Email_address	VARCHAR
Phone_number	VARCHAR
YTD_sales	DECIMAL
Salesrep_id	VARCHAR
Nationality	VARCHAR
National_id	VARCHAR

Open in SQL editor 🔗

Join keys Filters

After you select at least two columns of different data types, click **Preview** to ensure that the columns were properly joined. Preview can take a while if you are joining large tables. Click **Next** to continue joining these tables.

SALES_REP	CUSTOMERS
SALESREP_ID	SALESREP_ID

3. Click Continue.

Open SQL editor

You can run a CREATE VIEW SQL statement in the SQL editor to join your virtual objects. If you create the view in the SQL editor, do not return to the Join virtual objects page and retry your join in the user interface.

Back Continue

4. If any Query Customization is required, then we can change here and click on Run All button. For now, Click Back. Click Preview on the Previous screen to see the join data preview.

SQL editor

* Untitled - 1 Back

```

1 CREATE VIEW view_name
2   AS SELECT "ADMIN"."CUSTOMERS"."CUST_ID" AS "ADMIN_CUSTOMERS_CUST_ID", "ADMIN"."CUSTOMERS"."CUSTNAME" AS "ADMIN_CUSTOMERS_CUSTNAME",
3   "ADMIN"."CUSTOMERS"."CITY" AS "ADMIN_CUSTOMERS_CITY", "ADMIN"."CUSTOMERS"."STATE" AS "ADMIN_CUSTOMERS_STATE",
4   "ADMIN"."CUSTOMERS"."COUNTRY_CODE" AS "ADMIN_CUSTOMERS_COUNTRY_CODE", "ADMIN"."CUSTOMERS"."POSTAL_CODE" AS "ADMIN_CUSTOMERS_POSTAL_CODE",
5   "ADMIN"."CUSTOMERS"."EMAIL_ADDRESS" AS "ADMIN_CUSTOMERS_EMAIL_ADDRESS", "ADMIN"."CUSTOMERS"."PHONE_NUMBER" AS "ADMIN_CUSTOMERS_PHONE_NUMBER",
6   "ADMIN"."CUSTOMERS"."YTD_SALES" AS "ADMIN_CUSTOMERS_YTD_SALES", "ADMIN"."CUSTOMERS".SALESREP_ID AS "ADMIN_CUSTOMERS_SALESREP_ID",
7   "ADMIN"."CUSTOMERS"."NATIONALITY" AS "ADMIN_CUSTOMERS_NATIONALITY", "ADMIN"."CUSTOMERS".NATIONAL_ID AS "ADMIN_CUSTOMERS_NATIONAL_ID",
8   "ADMIN"."CUSTOMERS"."CREDITCARD_NUMBER" AS "ADMIN_CUSTOMERS_CREDITCARD_NUMBER", "ADMIN"."CUSTOMERS".CREDITCARD_TYPE AS "ADMIN_CUSTOMERS_CREDITCARD_TYPE",
9   "ADMIN"."CUSTOMERS"."CREDITCARD_EXP" AS "ADMIN_CUSTOMERS_CREDITCARD_EXP", "ADMIN"."CUSTOMERS".CREDITCARD_CVV AS "ADMIN_CUSTOMERS_CREDITCARD_CVV",
10  "ADMIN"."SALES REP"."FIRST_NAME" AS "ADMIN_SALES REP_FIRST_NAME", "ADMIN".SALES REP".LAST_NAME" AS "ADMIN_SALES REP_LAST_NAME",
11  "ADMIN".SALES REP".NATIONALITY" AS "ADMIN_SALES REP_NATIONALITY", "ADMIN".SALES REP".NATIONAL_ID" AS "ADMIN_SALES REP_NATIONAL_ID",
12  "ADMIN".SALES REP".PHONE_NUMBER" AS "ADMIN_SALES REP_PHONE_NUMBER", "ADMIN".SALES REP".AGE" AS "ADMIN_SALES REP_AGE",
13  "ADMIN".SALES REP".SEX" AS "ADMIN_SALES REP_SEX", "ADMIN".SALES REP".TERRITORY" AS "ADMIN_SALES REP_TERRITORY",
14  "ADMIN".SALES REP".EMAIL_ADDRESS" AS "ADMIN_SALES REP_EMAIL_ADDRESS", "ADMIN".SALES REP".MANAGER_ID" AS "ADMIN_SALES REP_MANAGER_ID"
15  FROM "ADMIN".CUSTOMERS", "ADMIN".SALES REP"
16  WHERE "ADMIN".CUSTOMERS".SALESREP_ID"="ADMIN".SALES REP".SALESREP_ID"

```

5. Close the Preview Popup.

The screenshot shows the 'Join virtual objects' step in the data management console. A tooltip is displayed over the 'TERRITORY' column in the preview table, indicating that multiple rows from the 'CUSTOMERS' table are being joined. The preview table contains data from both the 'SALES_REP' and 'CUSTOMERS' tables.

SALESREP_ID	FIRST_NAME	LAST_NAME	NATIONALITY	NATIONAL_ID	PHONE_NUMBER	AGE	SEX	TERRITORY
NC160	Betsy	Adams	UK	124-168-918	229-990-2162	35	F	NorthCentral
NC166	Penney	Hayes	FR	516-264-270	630-492-6535	50	F	NorthCentral
NC169	Fredericka	King	ES	766-887-613	448-788-1089	88	F	NorthCentral
NC169	Fredericka	King	ES	766-887-613	448-788-1089	88	F	NorthCentral
NC232	Lakesha	Jenkins	IT	910-870-499	828-268-4303	21	F	NorthCentral
NC232	Lakesha	Jenkins	IT	910-870-499	828-268-4303	21	F	NorthCentral

6. **Filters** tab can be used to filter some of the data from this view. For now leave it as default and click **Next**.

The screenshot shows the 'Join virtual objects' interface with the 'Filters' tab selected. It displays filter clauses for the 'CUSTOMERS' and 'SALES_REP' tables.

Column name	Data type
COUNTRY_CODE	VARCHAR
POSTAL_CODE	VARCHAR
EMAIL_ADDRESS	VARCHAR
PHONE_NUMBER	VARCHAR
YTD_SALES	DECIMAL
SALESREP_ID	VARCHAR
NATIONALITY	VARCHAR
NATIONAL_ID	VARCHAR
CREDITCARD_NUMBER	BIGINT

Column name	Data type
SALESREP_ID	CLOB
FIRST_NAME	CLOB
LAST_NAME	CLOB
NATIONALITY	CLOB
NATIONAL_ID	CLOB
PHONE_NUMBER	CLOB
AGE	CLOB
SEX	CLOB
TERRITORY	CLOB

Open in SQL editor

Join keys **Filters**

Programmatically filter data using simple clause statements like the one below.

```
"<SCHEMA>"<TABLE>"<COLUMN>" = <VALUE>
```

CUSTOMERS

```
"ADMIN"."CUSTOMERS"."CUST_ID" = <VALUE>
```

100 rows meet the filter criteria

SALES_REP

```
"ADMIN"."SALES_REP"."SALESREP_ID" = <VALUE>
```

7. Review the joined table, and click **Next**.

SALESREP_ID	FIRST_NAME	LAST_NAME	NATIONALITY	NATIONAL_ID	PHONE_NUMBER	AGE
NC160	Betsy	Adams	UK	124-168-918	229-990-2162	35
NC166	Penney	Hayes	FR	516-264-270	630-492-6535	50
NC169	Fredericka	King	ES	766-887-613	448-788-1089	88
NC169	Fredericka	King	ES	766-887-613	448-788-1089	88
NC232	Lakeesha	Jenkins	IT	910-870-499	828-268-4303	21
NC232	Lakeesha	Jenkins	IT	910-870-499	828-268-4303	21

8. For the view name, type joined_customers_sales_table. Select a project from the list that was created earlier, eg. Data Virtualization Project. Click **Create view**.

Assign to (all tables will be assigned to the same project)

Data request (1) Project (1) Virtualized data

Data Virtualization Project

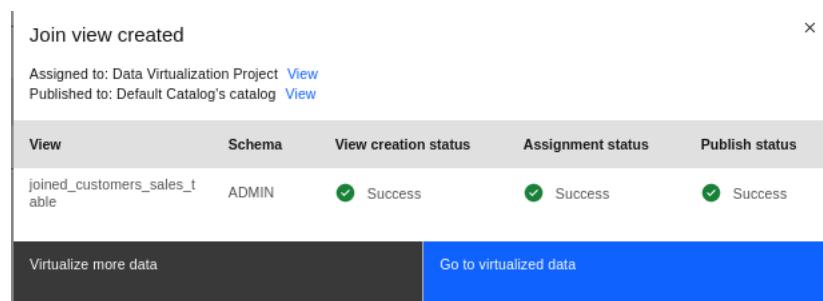
Publish to

Catalog (1)

Default Catalog

View name	Schema name
joined_customers_sales_table	ADMIN

9. When the process completes, Click **Go to Virtualized Data**.



10. Navigate to Projects -> Data Virtualization Project to view the project to preview the virtualized data.

Table	Schema name	Created on	Statistics last collected on
joined_customers_sales_table	ADMIN	Sep 2, 2022 11:28 AM	Not applicable
SALESREP	ADMIN	Sep 2, 2022 11:21 AM	Not collected
CUSTOMERS	ADMIN	Sep 2, 2022 11:21 AM	Not collected

11.Click on the project created earlier.

Name	Date created	Your role	Collaborators
Data Virtualization Project	3 days ago	Admin	AA
Banking Demo Project	2 months ago	Admin	AA
Datastage Demo Assets	2 months ago	Admin	AA

12.Navigate to **Assets** Tab. You will see the virtual view and the connect object to the Data Virtualization Service.

13.Normally, You will need a Credential or API key to view the data in the project. This credential can be saved in the connection object so that it's not prompted again and again. Click on the Connection object eg. DSXXXXXX.

Name	Last modified
ADMIN.joined_customers_sales_table	4 minutes ago admin (You)
DS16621325392290910	4 minutes ago admin (You)

14.Click on **Credentials** in the left navigation pane menu.

Review the connection information.

Connection overview

Name: DS16621325392290910

Description: Connection description

Connection details

Database*: BIGSQL

Cancel Save

15. Select one of the Authentication Method as the personal connection.

Review the connection information.

Credentials

Credential setting: Personal

Authentication method*

Port is SSL-enabled

SSL certificate

```
-----BEGIN CERTIFICATE-----
MIIDBzCCAfegAwIBAgIRAljyOHEkIpGPu/kYybPFWFUwDQYJKoZIhvNAQELBQA
wHTEBMBkGA1UEAxMSemVuLWNhLNicnRpZmljYXRIMBAXDjTyMDYzMDiMDQyN1oX
DTI1MDY0TlxMDkyN1owHTEBMBkGA1UEAxMSemVuLWNhLNicnRpZmljYXRIMIIB
IANBeknnkiG9wvRA0FFAAOCAB8AMTBCeKCAQFA0TOY5Nkuw3s9H66lII/vCfrx
-----END CERTIFICATE-----
```

Cancel Save

16. Chose User Name and Password.

Test connection

Connection overview
Connection details
Credentials
Certificates

Credentials
Credential setting ⓘ
Personal Shared

Authentication method* ⓘ
...
API key
Username and password
SSL certificate ⓘ
-----BEGIN CERTIFICATE-----
MIIDBzCCAf+gAwIBAgIRAIyOHEkIpGPu/kYybPFWFUwDQYJKoZIhvNaQELBQAw
HTEBMkGAUExMSevVulWNlWNIcnRpZmlnYXRIMB4DTjMDY2MDixMDQyN1oX
DTI1MDYyOTIxMDkyN1owHTEBMkGA1UEAxMSemvVulWNlWNIcnRpZmlnYXRIMB
TIANRskahkiG9w0BA0FFAAOCAO8AMTICeKCAOEAOTy5Nkiu3s9H66Jii/cIfTx

Cancel Save

17. Enter the CP4D credentials and click on **Test Connection**. It should be successful. Click **Save** to save the credential information in the connection.

Test connection

Connection overview
Connection details
Credentials
Certificates

Credentials
Credential setting ⓘ
Personal Shared

Authentication method* ⓘ
Username and password

Username* ⓘ
admin

Password* ⓘ

Certificates
Port is SSL-enabled ⓘ

SSL certificate ⓘ

Cancel Save

The test was successful.
Click Save to update the connection information.

18. Now Click on the Virtualized Data Asset ADMIN.joined_customers_sales_table.

Projects / Data Virtualization Project

Overview Assets Jobs Manage

Find assets

2 assets

All assets

Name	Last modified
ADMIN.joined_customers_sales_table	11 minutes ago admin (You)
DS16621325392290910	11 minutes ago admin (You)

Add asset New asset

Asset types
Data access
Data

Data in this project
Drop data files here or browse for files to upload

19. You should be able to preview the data without being prompted for credentials.

The screenshot shows the 'Preview' tab selected in the top navigation bar. The main area displays a table with 26 columns and 10 rows of data. The columns include 'SALESRE...', 'FIRST_N...', 'LAST_N...', 'ADMIN_SALESREP_NATIONA...', 'ADMIN_SALESREP_NATIONA...', 'ADMIN_SALESPHONE_NU...', 'AGE', and 'SEX'. The data shows various names, countries, phone numbers, and ages. A 'Refine' button is located in the top right corner of the preview area.

20. Optionally you can click on **Refine** button to explore or refine this data using **Watson Data Refinery**. The part of virtualized data can be saved as CSV file in the project using the Data Refinery Jobs.

Next steps

Now your virtual data is ready to be used. For example, you can these tasks:

Analyze the data in a notebook

- 1) Under the Data Virtualization Project, In **Assets** Tab, Click on **New Asset +** to add a notebook Asset.

The screenshot shows the 'Assets' tab selected in the top navigation bar. On the left, there's a sidebar with 'Asset types' and counts for 'Data access' (1) and 'Data' (1). The main area lists 'All assets' with two entries: 'ADMIN.joined_customers_sales_table' (modified 11 minutes ago) and 'DS16621325392290910' (Connection, modified 11 minutes ago). A 'New asset +' button is highlighted in blue at the top right of the asset list.

- 2) Click on **Jupyter notebook Editor**.

New asset

Select the tool to create an operational or configuration asset.

The screenshot shows the 'New asset' interface with a search bar at the top. On the left, a sidebar lists 'Tool type' categories: All types, Automated builders, Graphical builders, Code editors, Component editors, and Data access tools. The 'All types' category is selected. The main area displays three tool categories: Pipelines, Code editors, and Component editors. Under Pipelines, the 'SPSS Modeler' tool is shown with a brief description. Under Code editors, 'Federated Learning' and 'Jupyter notebook editor' are listed. Under Component editors, 'DataStage component' and 'Parameter set' are shown. A toggle switch at the bottom left is set to 'Show descriptions'.

3) Enter the Notebook Name and Credentials and Click **Create.**

The screenshot shows the 'New notebook' dialog. It has tabs for 'Blank', 'From file', and 'From URL', with 'Blank' selected. The 'Name' field contains 'Data Analysis Notebook'. The 'Select runtime' dropdown is set to 'IBM Runtime 22.1 on Python 3.9 (1 vCPU, 2 GB RAM)'. Below it, a note says 'The selected runtime has 1 vCPU and 2 GB RAM.' The 'Language' section shows 'Python 3.9' selected. At the bottom right are 'Cancel' and 'Create' buttons, with 'Create' highlighted in blue.

4) Click on icon to View the Find and Add Data.

The screenshot shows the Jupyter Notebook interface. The top navigation bar includes 'File', 'Edit', 'View', 'Insert', 'Cell', 'Kernel', 'Widgets', and 'Help'. The status bar indicates 'Not Trusted | Python 3.9' and 'CPU: Memory:218.1 MB / 2 GB'. The main area shows a code cell starting with 'In []:'.

5) Under the Files tab, Click **Insert to Code -> pandas Data Frame For the Virtual Table in the project.**

- 6) The code will be added automatically. Click Run button in the toolbar to execute the code.

```

In [ ]: import itc_utils.flight_service as itcfs
readClient = itcfs.get_flight_client()

# NOTE:
# A limit of 5000 rows has been applied to the request to enable sample previewing.
# Adjust the display message as needed by editing the following lines:
from IPython.core.display import display, HTML
display(HTML('A row limit of 5000 has been applied to the query to enable sample previewing. If the data set is larger, only the first 5000 row
# Edit select_statement to change or disable the row limit.
DS16621325392290910 data_request = {
    'connected_data_name': """ADMIN.joined_customers_sales_table""",
    'interaction_properties': {
        'row_limit': 5000
    }
}

flightInfo = itcfs.get_flight_info(readClient, nb_data_request=DS16621325392290910_data_request)

data_df_1 = itcfs.read_pandas_and_concat(readClient, flightInfo)
data_df_1.head(10)

```

- 7) You should be able to access the virtualized data in the Watson Studio Notebook.

SALESPER_ID	FIRST_NAME	LAST_NAME	ADMIN.SALES.REP.NATIONALITY	ADMIN.SALES.REP.NATIONAL_ID	ADMIN.SALES.REP.PHONE_NUMBER	AGE	SEX	TERRITORY	AL	
0	NC160	Betsy	Adams	UK	124-168-918	229-990-2162	35	F	NorthCentral	Bi
1	NC166	Penney	Hayes	FR	516-264-270	630-492-6535	50	F	NorthCentral	Per
2	NC169	Fredericka	King	ES	766-887-613	448-788-1089	88	F	NorthCentral	Fred
3	NC169	Fredericka	King	ES	766-887-613	448-788-1089	88	F	NorthCentral	Fred
4	NC232	Lakesha	Jenkins	IT	910-870-499	828-268-4303	21	F	NorthCentral	Lakeer
5	NC232	Lakesha	Jenkins	IT	910-870-499	828-268-4303	21	F	NorthCentral	Lakeer

Visualize the data with a dashboard

- 1) Under the Data Virtualization Project, In **Assets** Tab, Click on **New Asset +** to add a new Dashboard.

The screenshot shows the 'Assets' tab selected in the top navigation bar. On the left, there's a sidebar with sections for 'Overview', 'Assets' (selected), 'Jobs', and 'Manage'. A search bar at the top right says 'Search'. To the right, a large table lists 'All assets' with columns for 'Name' and 'Last modified'. Three assets are listed: 'Data Analysis Notebook' (Notebook), 'ADMIN.joined_customers_sales_table' (application/octet-stream), and 'DS16621325392290910' (Connection). A 'Data in this project' section on the right has a placeholder 'Drop data files here or browse for files to upload'.

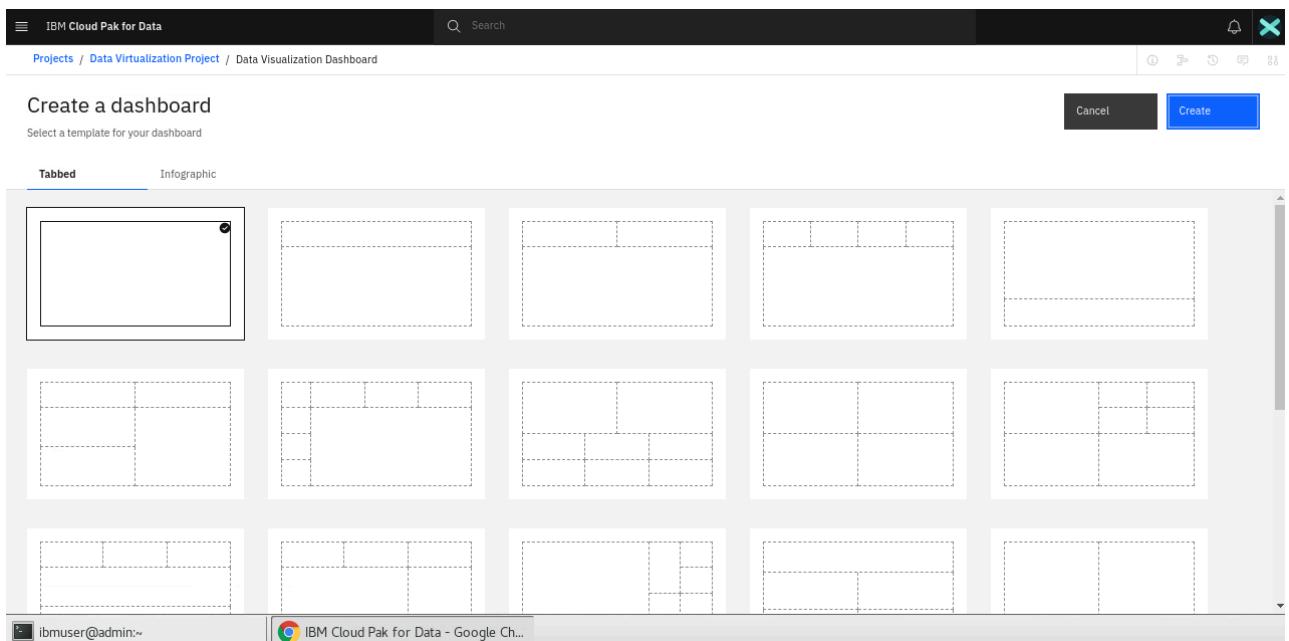
2) Click on Dashboard Editor.

The screenshot shows the 'New asset' dialog. The 'Tool type' dropdown is set to 'All types'. Below it, there are sections for 'Automated builders' (AutoAI, Metadata enrichment) and 'Graphical builders' (Dashboard editor, Data Refinery, DataStage, Decision Optimization). Each tool has a brief description and a small icon. At the bottom left is a 'Show descriptions' checkbox, and at the bottom right are 'Cancel' and 'Create' buttons.

3) Enter the name and description and click Create.

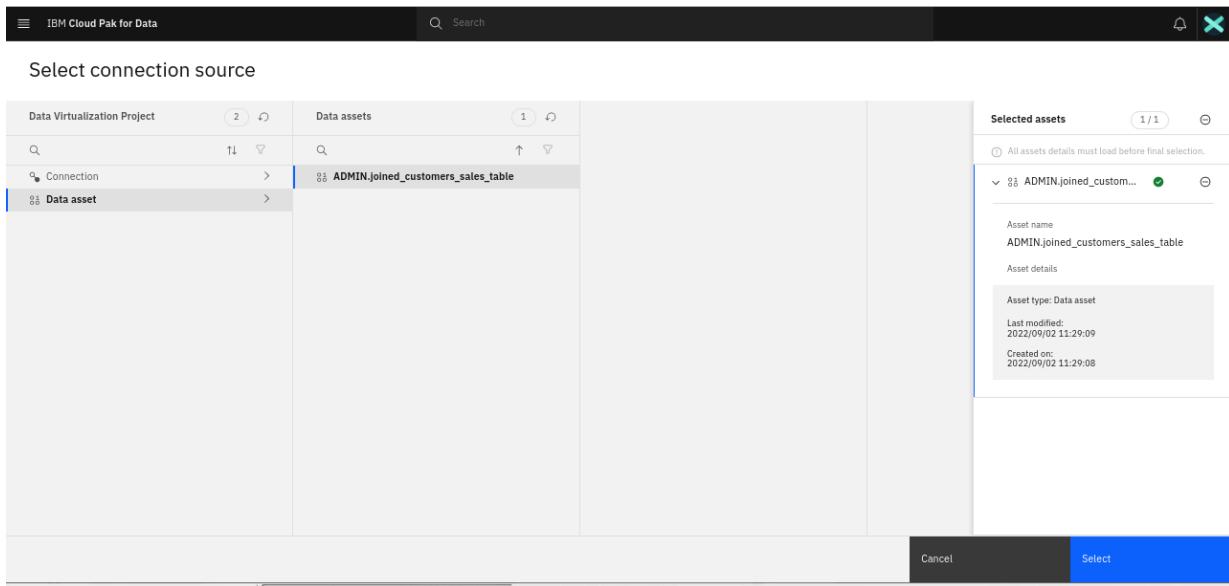
The screenshot shows the 'Create a dashboard' dialog. In the 'Define details' section, the 'Name' field is filled with 'Data Visualization Dashboard'. The 'Description (Optional)' field contains 'Data Visualization Dashboard'. At the bottom right are 'Cancel' and 'Create' buttons, with 'Create' being the active one.

4) Click Create to create an empty tabbed Dashboard.



- 5) Click **Select a source +** button in the Data Pane to select a data source.

- 6) Select the Virtual Table. Click **Select**.

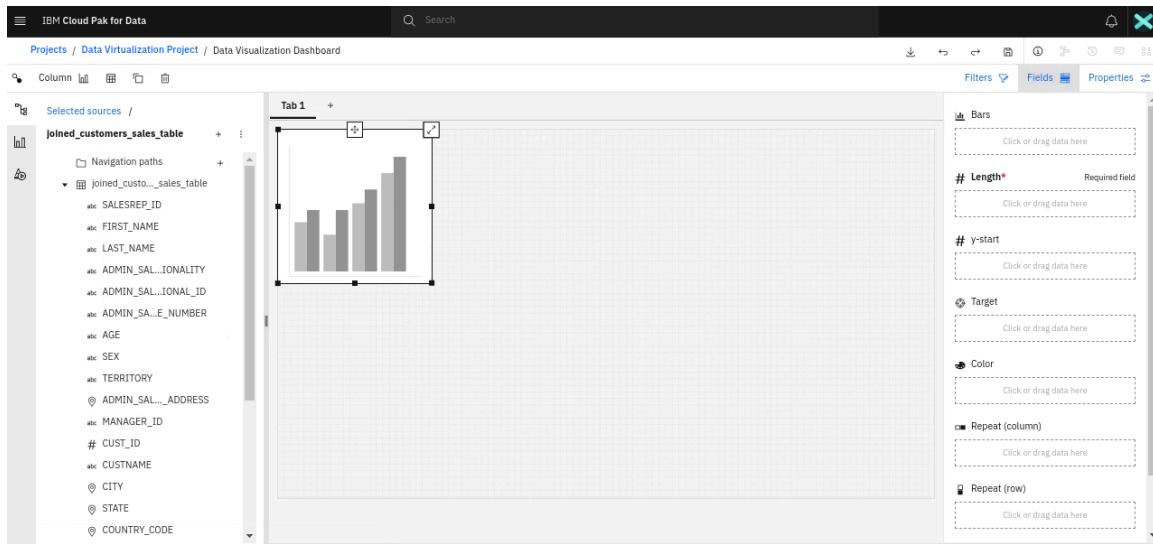


- 7) The Data source will be added successfully. Click on **Visualization Tab**  to see the types of Visualization to be used.

- 8) Click on the required visualization to add it to the dashboard.

9) Configure the chart to show the required visualization. Eg. X Axis/Y Axis fields.

Alternatively, New charts can also be created by directly dragging and dropping a field from the virtual table to the dashboard.



Appendix

Below steps are for reference only.

DB2 Instance Provisioning

1. If you have a *Db2* service listed, then there is no need to provision another instance. Otherwise, follow these steps to procure a DB2 instance.
2. Under the services selection page, select **DB2**. Note that the service status is Available. For rest, the status is enabled.

Service	Status	Description
Data Virtualization	Enabled	Query many data sources as one.
Db2	Available	Relational database that delivers advanced data management and analytics capabilities for transactional and warehousing workloads.
RStudio Server with R3.6	Enabled	Optional development environment for

3. Click **New Instance**.

Type: Database
Version: 11.5.7.0-cn5-x86_64
Provider: IBM
Category: Data sources
Related links: Docs

Summary

Relational database that delivers advanced data management and analytics capabilities for transactional and warehousing workloads.

SQL

SQL	Number of executions	Statement execution time	Estimated runtime	CPU time (ms)	Rows read	Rows returned	Rows modified	Coordinator State
<code>call get_device_info@DB2CAST (7 AS INT)</code>	5	0.69.013	--	0.00.001	0	1	0	
<code>CALL SYSPROC.SYSPINSTALLOBJECTCAST L...</code>	1	0.69.004	--	0.000	0	0	0	
<code>with form as (select distinct host_name, MEM...</code>	5	0.69.002	--	0.00.002	0	1	0	
<code>LOCK TABLE SYSTEM STMO, DB2DE_INF...</code>	5	0.69.002	--	0.00.002	0	0	0	
<code>select CPU_SYSTEM as cpusystem, CPU_US...</code>	10	0.69.001	--	0.00.001	0	1	0	
<code>SELECT TRIGNAME FROM SYSCAT.TRIGGER...</code>	1	0.69.001	--	0.000	4	4	0	

4. Provide the DB Details like DB Name (Max 8 Chars). Scroll Down.

Create a database

Configure

Database name: Customer

Number of nodes: 1

CPU per node for Db2: 2.1

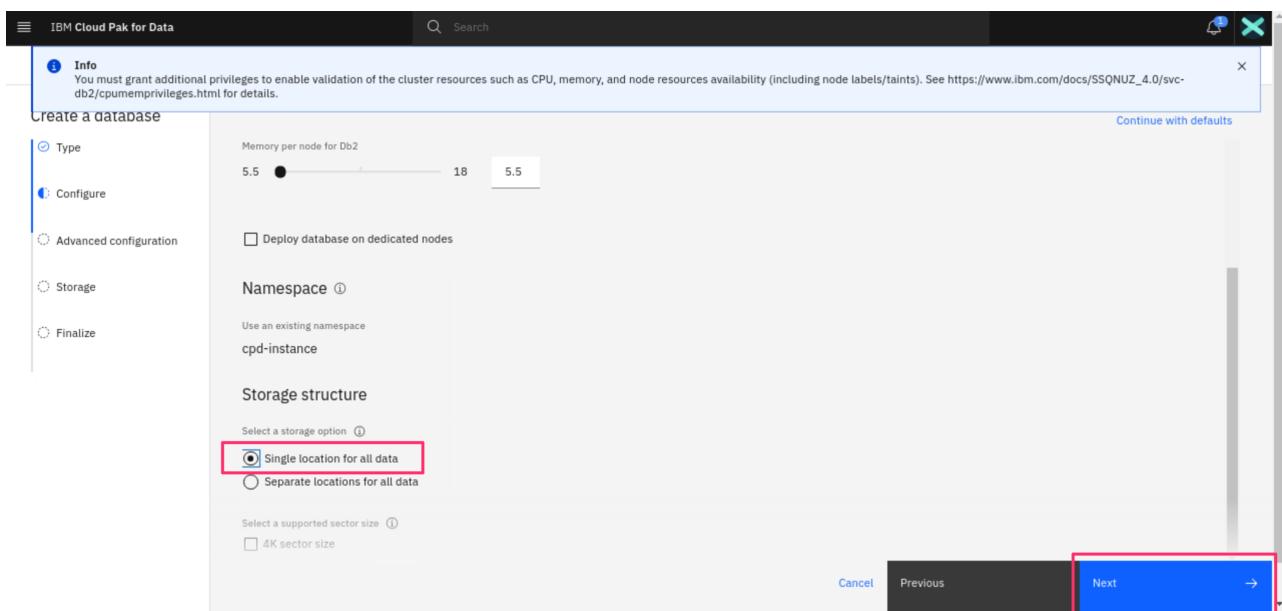
Memory per node for Db2: 5.5

Deploy database on dedicated nodes:

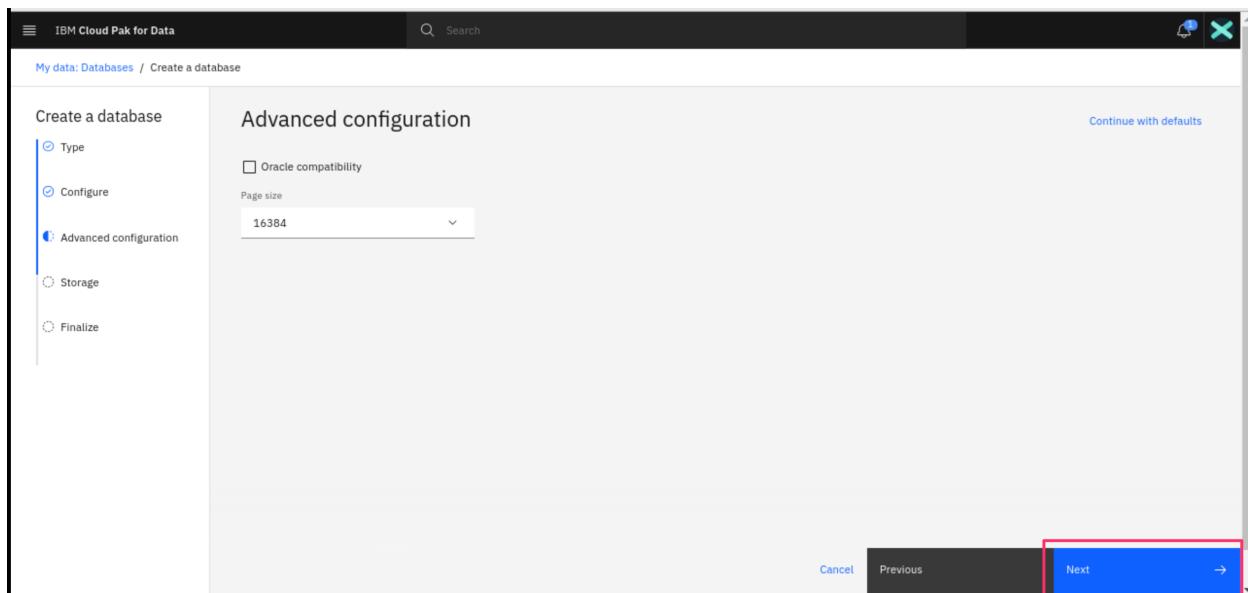
Namespace: 1

Continue with defaults

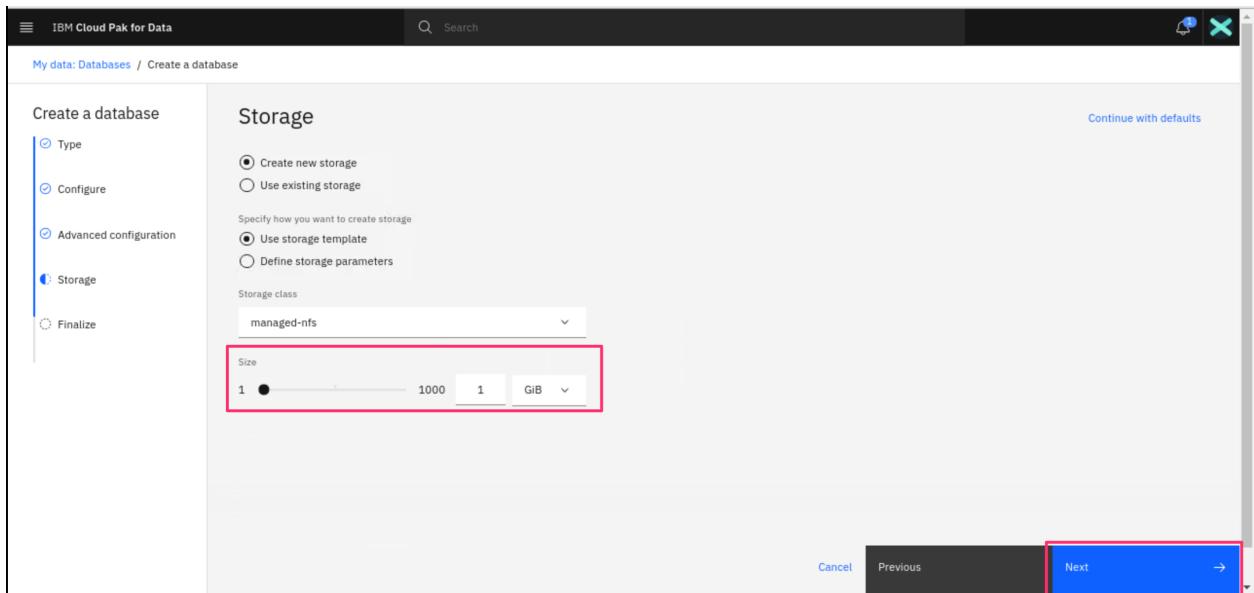
5. Select **Single Location for all Data** Option and click **Next**.



6. Click **Next** on the advanced configuration.



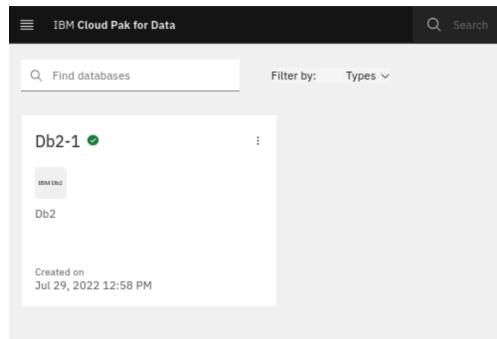
7. Configure the Storage Size as **10 GiB** and Click **Next**.



8. Click **Create** on the summary screen.

9. It will start provisioning the service.

10. The service will be provisioned in 10-20 min depends on the configuration chosen.



11. Verify that the services are provisioned on your *Service instances* page.

Instances								
Filter by: Type ▾ Status ▾ Find instances								
Name	Type	Created by	vCPU requests	Memory requests (GiB)	Users	Status	Created on	
Db2-1 Service instance for db2oltp-16591138850024...	db2oltp	admin	2.20	5.75 Gi	1	Green	Jul 29, 2022	⋮
ds-px-default The default DataStage runtime instance	datastage	admin	2.50	6.00 Gi	1	Green	Jul 12, 2022	⋮
data-virtualization	dv	admin	11.50	38.50 Gi	1	Green	Jul 6, 2022	⋮
data-management-console Data Management Console	dmc	admin	4.20	10.62 Gi	1	Green	Jul 6, 2022	⋮
openscale-defaultinstance IBM Watson OpenScale	aios	admin	0.00	0.00 Gi	1	Green	Jul 1, 2022	⋮