

# Db2 on Cloud and Watson Studio

## Workshop Guide

### Version 3

Analyzing airline flight delays

*SQLAdria & PDUG 2023*

*June 15 and 16, 2023*

*Roger E. Sanders*

*Principal Sales Enablement and Skills Content Specialist - Data and AI*

[roger.sanders@us.ibm.com](mailto:roger.sanders@us.ibm.com)

## NOTICES AND DISCLAIMERS

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. Information regarding potential future products is intended to outline potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract.

The development, release, and timing of any future features or functionality described for our products remains at our sole discretion I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results like those stated here.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. **This document is distributed "as is" without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.** IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply."

**Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.**

Performance data contained herein was generally obtained in controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions, and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to ensure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products about this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. **IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a purpose.**

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates. Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

OpenShift is a trademark of Red Hat, Inc.

UNIX is a registered trademark of The Open Group in the United States and other countries.

© 2023 International Business Machines Corporation. No part of this document may be reproduced or transmitted in any form without written permission from IBM.

**U.S. Government Users Restricted Rights — use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.**

# Table of Contents

<b>I. Introduction .....</b>	<b>5</b>
The Business Use Case for this workshop .....	7
<b>II. Provisioning a Db2 on Cloud Lite plan database on the IBM Cloud .....</b>	<b>9</b>
1. Create an IBM account. ....	9
2. Create an IBM Cloud account. ....	14
3. Provision a Db2 on Cloud Lite plan database.....	29
4. Create a service credential for the Db2 on Cloud database. ....	35
5. Query the system catalog using the Db2 on Cloud User Interface (UI). ....	38
<b>III. Building and populating a Db2 on Cloud database table .....</b>	<b>47</b>
1. Create a table named AIRLINE_DELAY_CAUSE. ....	47
2. Populate the AIRLINE_DELAY_CAUSE table. ....	54
<b>IV. Using IBM Watson Studio to run a Jupyter Notebook .....</b>	<b>64</b>
1. Create an IBM Cloud Pak for Data as a Service account. ....	64
2. Create a new project. ....	68
3. Add a Jupyter Notebook to the project. ....	71
4. Add a data file to the project. ....	76
5. Prepare the Jupyter Notebook for execution. ....	79
6. Run the Jupyter Notebook.....	87
<b>V. Analyzing the data.....</b>	<b>90</b>
1. Log in to your IBM Cloud Pak for Data as a Service account. ....	90
2. Add a new Jupyter Notebook to the project created earlier. ....	92
3. Prepare the Jupyter Notebook for execution.....	98
4. Run the Airline_Delay_Analysis Jupyter Notebook.....	101
<b>VI. Thank you! .....</b>	<b>107</b>

## Acknowledgements

Special thanks to **George Baklarz** for his assistance with the **Db2 Magic Commands** and his **Jupyter Notebook** debugging expertise. Without George's help, the workshop exercise would have taken much longer to complete.

And thanks to Danny Arnold for putting together the first two versions of this workshop guide.

# I. Introduction

As the COVID-19 pandemic winds down and life returns to normal, the return to business and leisure travel has resulted in many challenges for travelers – most notably, an apparent increase in flight delays and cancellations. However, because many people have spent the last 2-3 years *not* traveling, have flight delays and cancellations really increased over the past year (2022)? Or is this a historical pattern that people have simply forgotten about? This workshop attempts to answer that.

To determine whether flight delays and cancellations were up in 2022, data containing airline on-time statistics and delay causes for the period of January 2019 through December 2022 was downloaded from the **United States Department of Transportation, Bureau of Transportation Statistics** website ([https://www.transtats.bts.gov/ot\\_delay/](https://www.transtats.bts.gov/ot_delay/)). This data contains information about every flight, flight delay, flight cancellation, and reason for the delay or cancellation (there can be multiple reasons that contribute to a flight delay or cancellation) for every commercial airport and airline operating in the United States during this time.

This workshop is designed to provide a person with hands-on experience in provisioning a Db2 on Cloud Lite (free) plan database on the IBM Cloud, creating an AIRLINE\_DELAY\_CAUSE table in the Db2 on Cloud database, loading that table with data obtained from the United States Department of Transportation, Bureau of Transportation Statistics website, and then analyzing the data using IBM Watson Studio in IBM Cloud Pak for Data as a Service and Jupyter Notebook.

This workshop is divided into four main sections, as follows:

## 1. Provisioning a Db2 on Cloud Lite plan database on the IBM Cloud.

- Create an IBM account and IBMid (*optional – not necessary if you already have a valid IBMid*).
- Create an IBM Cloud account (*optional – not necessary if you already have an IBM Cloud account*).
- Provision a Db2 on Cloud Lite (free) plan database in the IBM Cloud data center of your choice.  
**IMPORTANT:** The Db2 on Cloud Lite plan is only available in the Dallas (us-south) and London (eu-gb) data centers.
- Create a service credential for the newly created Db2 on Cloud database that will be used to perform some of the exercises in this workshop.
- Query the system catalog of the newly created Db2 on Cloud database using the Db2 on Cloud User Interface.

## 2. Building and populating a Db2 on Cloud database table.

- Create a table named **AIRLINE\_DELAY\_CAUSE** using the Data Definition Language (DDL) statement stored in a file named **AIRLINE\_DELAY\_CAUSE\_DDL.sql** and the Db2 on Cloud User Interface.
- Populate the **AIRLINE\_DELAY\_CAUSE** table with data stored in a file named **Airline\_Delay\_Cause.csv** using the Db2 on Cloud User Interface.

### 3. Using IBM Watson Studio to run a Jupyter Notebook

- Create an IBM Cloud Pak for Data as a Service account (*optional – not necessary if you already have an IBM Cloud Pak for Data as a Service account.*)
- Create a new project in Watson Studio/ IBM Cloud Pak for Data.
- Add a Jupyter Notebook (named **Build\_And\_Populate\_Table**) to the project by importing a pre-built Jupyter Notebook file.
- Add a data file to the project.
- Prepare the Jupyter Notebook for execution by:
  - Adding information found in the Db2 on Cloud service credential created earlier to application variables that are used to establish a database connection.
  - Generating a code snippet that will be used to access the data file that was added to the project.
- Run the **Build\_And\_Populate\_Table** Jupyter Notebook.

### 4. Analyzing the data

- Log in to your IBM Cloud Pak for Data as a Service account.
- Add another Jupyter Notebook (named **Analyze\_Delay\_Cause**) to the project by importing another pre-built Jupyter Notebook file.
- Prepare the new Jupyter Notebook for execution by adding information found in the Db2 on Cloud service credential created earlier to application variables that are used to establish a database connection.
- Run the **Analyze\_Delay\_Cause** Jupyter Notebook.

#### The data set used for this workshop

Here is some background information about the data set that is used in this workshop:

- The data used for this workshop is data that has been collected by the *United States Department of Transportation* and consists of information about domestic (US) airline flight delays for the years 2013 through 2022. Information that can be gleaned from this data include the number of on-time flight departures and the number of flights that have been canceled or delayed. According to the definition provided by the Bureau of Transportation Statistics for the data set, a flight is considered delayed if it arrives at its destination airport 15 minutes or more past its scheduled arrival time. For delayed flights, the reason for the delay is provided from several categories:
  - Air carrier delay – This category covers flight delays that occur because of mechanical issues, personnel issues (such as a flight crew exceeding their hours), and other related reasons.
  - Aircraft arriving late – This category covers flight delays caused by the incoming flight (that is providing the aircraft for the flight) arriving late from its origin airport.
  - Security delay – This category covers flight delays that are caused by airport security reasons. For example, if a traveler gets past a Transportation Security Administration (TSA) checkpoint after a prohibited item is found, thereby breaching airport security, and causing all passengers to be re-screened.
  - National Aviation System (NAS) delay – This category covers flight delays that are caused by NAS issues, such as air traffic control and related issues.
  - Weather delay – This category covers flight delays that occur because of extreme weather such as thunderstorms, wintery weather and icing conditions, tornado warnings, and other weather-related delays.

### The Business Use Case for this workshop

A fictional travel company, Global Travel Associates (GTA), has been receiving negative scores on Customer Satisfaction Surveys that are completed by their customers traveling in the United States (U.S.). Their number one complaint is about flight delays that result in missed connections or late arrivals to their destinations.

To try to improve their customer satisfaction scores, GTA has obtained recent flight delay data from the U.S. Department of Transportation's Bureau of Transportation Statistics for the years 2019 through 2022. Their plan is to analyze this data in an effort to determine if flight delays have increased over the past year (2022) and if so, what might be causing the delays. To save on costs, GTA has selected the **Db2 on Cloud ‘Lite’ plan** as their cloud database platform. They also have decided to use the free tier of **IBM Watson Studio in IBM Cloud Pak for Data as a Service**.

The exercises in this workshop are designed to take you through GTA’s process of storing data obtained from U.S. Department of Transportation’s Bureau of Transportation Statistics for the years 2019 through 2022 in a Db2 on Cloud ‘Lite’ plan database and analyzing the data to gain insight on what has been causing flight delays and how flight delays have been trending. You will also see how GTA used a layered approach in examining the data to avoid making an incorrect assumption based on initial analysis results. Finally, you will see how GTA used historical flight delay data (obtained from the same source for years 2013 through 2022) to see how flight delays observed in 2022 compared to flight delays seen in the past.

At the conclusion of the workshop exercise, the participant should know how to provision a Db2 on Cloud database, create and populate a database table using both the Db2 on Cloud User Interface and source code stored in a Jupyter Notebook (executing in IBM Watson Studio on IBM Cloud Pak for Data as Service), and how to use a Jupyter Notebook and Db2 Magic Commands to analyze data quickly.

---

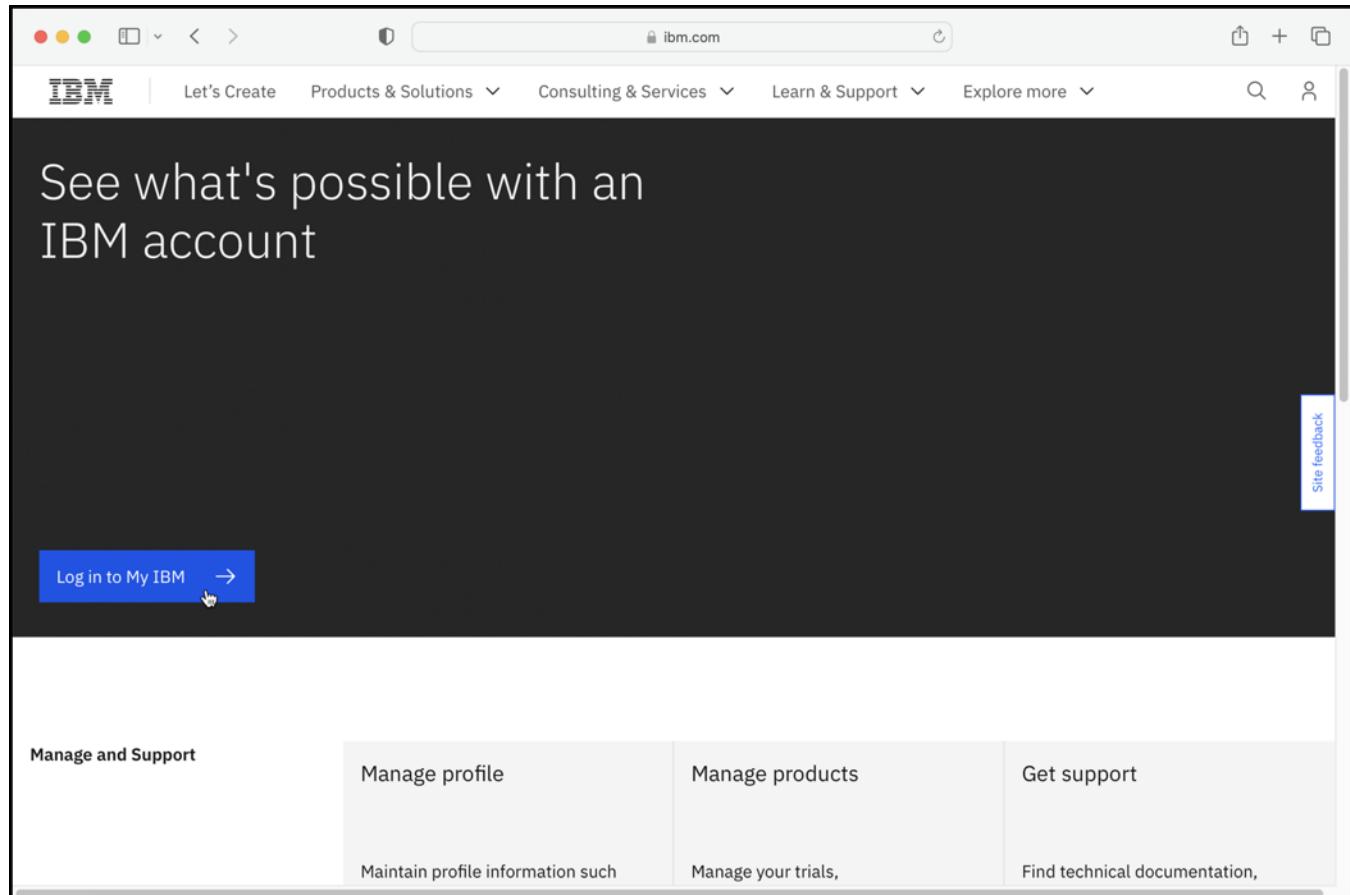
## II. Provisioning a Db2 on Cloud Lite plan database on the IBM Cloud

The premise behind this workshop is that a client wants to store and analyze flight delay information at the lowest possible cost. Consequently, a Db2 on Cloud Lite plan database is used since this offering is available at no charge.

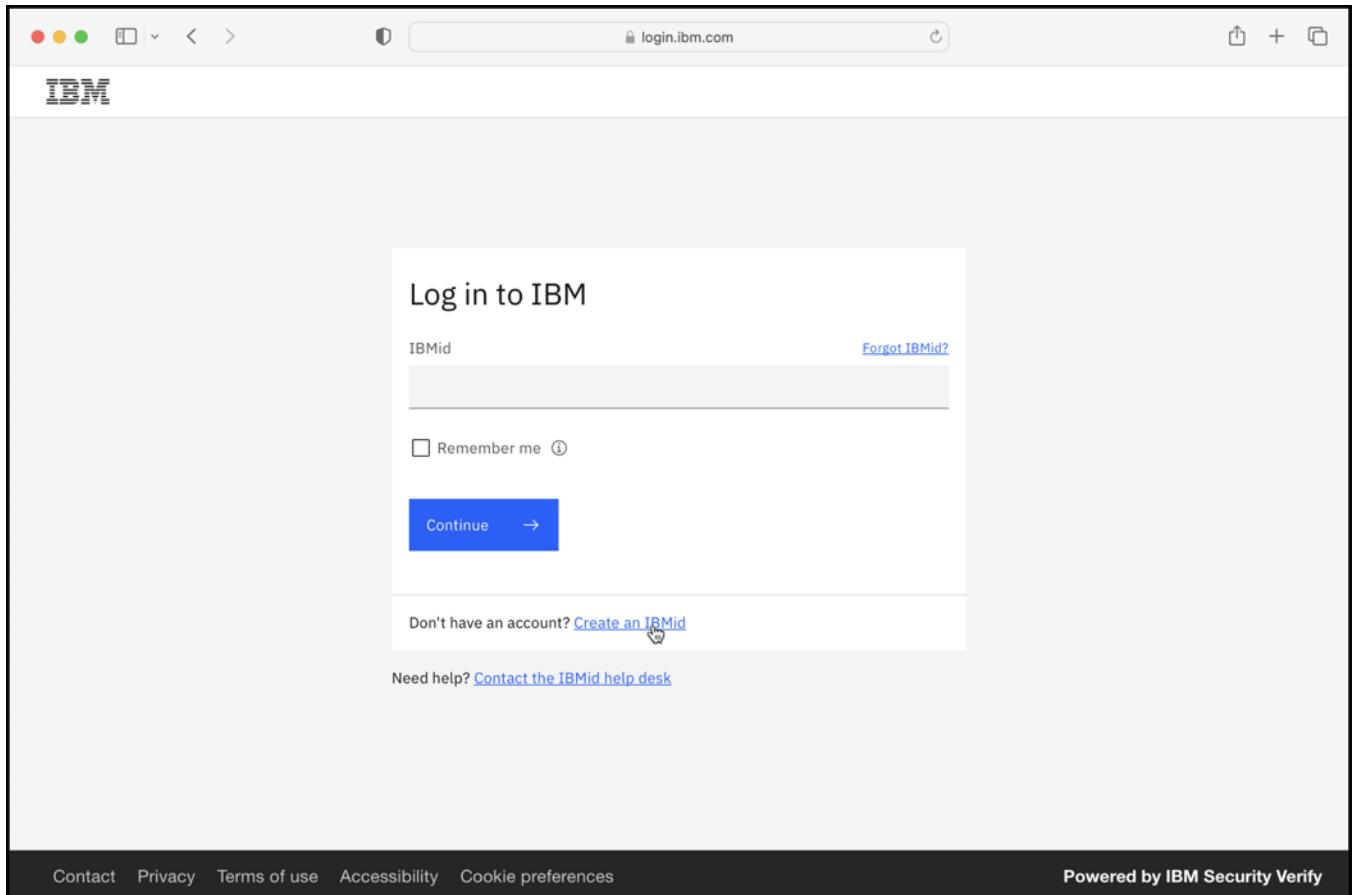
In order to provision a Db2 on Cloud “Lite” plan database, you must have an IBMid and an IBM Cloud account, both of which can be created free of charge. **If you already have an IBMid and an IBM Cloud account, proceed directly to Step 3.**

### 1. Create an IBM account.

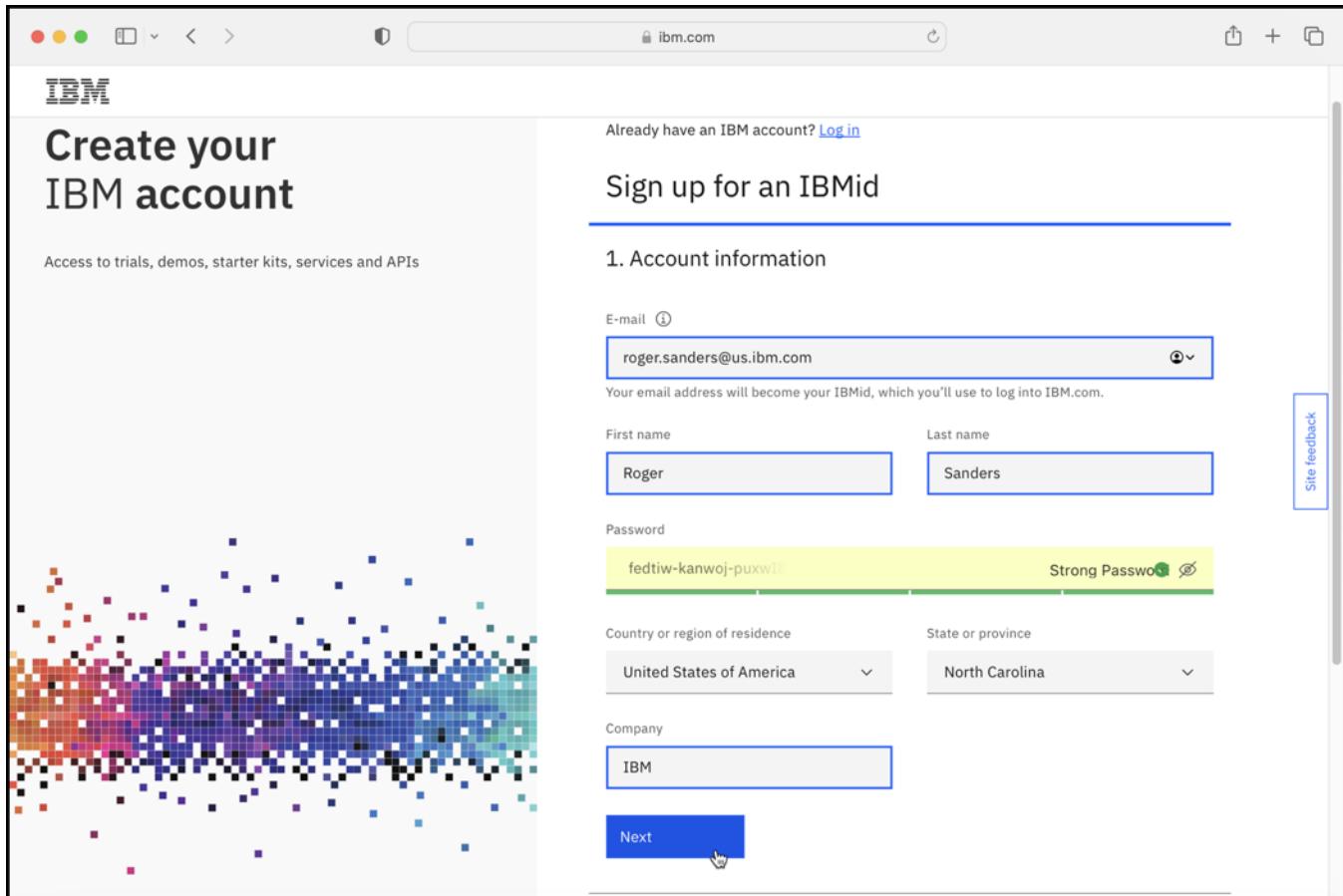
\_\_\_\_ 1. Open a web browser and go to the *IBM Account* web site (<https://www.ibm.com/account/us-en/>). Once there, select the **Log in to My IBM** button.



\_\_\_\_ 2. When the *Log in to IBM* popup is displayed, find the **Create an IBMid** link located at the bottom of the box and click on it.



\_\_\_\_ 3. When the *Create your IBM account* screen is displayed, fill out the form provided with the information requested. When finished, click the **Next** button.



The screenshot shows the 'Create your IBM account' sign-up page on ibm.com. The page features a large, colorful abstract graphic on the left. At the top right, there's a 'Sign up for an IBMid' button. Below it, a section titled '1. Account information' contains fields for E-mail (roger.sanders@us.ibm.com), First name (Roger), Last name (Sanders), Password (fedtiw-kanwoj-puxw123), Country or region of residence (United States of America), State or province (North Carolina), and Company (IBM). A 'Next' button is at the bottom. A 'Site feedback' link is visible on the right.

IBM

Create your  
IBM account

Access to trials, demos, starter kits, services and APIs

Already have an IBM account? [Log in](#)

Sign up for an IBMid

1. Account information

E-mail ⓘ

roger.sanders@us.ibm.com  ⓘ ↗

Your email address will become your IBMid, which you'll use to log into IBM.com.

First name

Roger

Last name

Sanders

Password

fedtiw-kanwoj-puxw123  ⓘ ↗

Strong Password

Country or region of residence

United States of America

State or province

North Carolina

Company

IBM

Next

Site feedback

4. The information you entered will then be displayed and a 7-digit confirmation code will be sent to the email address you provided. (The email will come from *IBM Security* and the subject will be *New User Registration*.) Once you receive an email with a confirmation code, enter the code in the **Verification token** field. (You may have to scroll down to see this field.) Then, locate and click the **Create account** button shown at the bottom of the page. (*Note: The confirmation code is only valid for 30 minutes. If, after 5 minutes, you don't see an email containing the code in your inbox, check your spam/junk email folder.*)

IBM

## Create your IBM account

Access to trials, demos, starter kits, services and APIs

Already have an IBM account? [Log in](#)

### Sign up for an IBMid

---

**1. Account information** [Edit](#)

E-mail	roger.sanders@us.ibm.com
First name	Roger
Last name	Sanders
Country	United States of America
State or province	North Carolina
Company	IBM

---

**2. Verify email**

We emailed a 7 digit code to  
[roger.sanders@us.ibm.com](mailto:roger.sanders@us.ibm.com)

This code will expire in 30 minutes.

Verification token

7575173

Didn't receive the email? Check your spam filter for an email from [ibmacct@iam.ibm.com](mailto:ibmacct@iam.ibm.com).

[Resend code](#)

---

IBM may use my contact data to keep me informed of products, services and offerings:

by email

You can withdraw your marketing consent at any time by submitting an [opt-out request](#). Also you may unsubscribe from receiving marketing emails by clicking the unsubscribe link in each email.

More information on our processing can be found in the [IBM Privacy Statement](#). By submitting this form, I acknowledge that I have read and understand the IBM Privacy Statement.

I accept the product [Terms and Conditions](#) of this registration form.

[Create account](#)

---

Contact IBM   Privacy   Terms of use   Accessibility   [Cookie Preferences](#)

\_\_\_\_ 5. When the *About your IBMid Account* popup appears, read through the terms and conditions of your IBM account. When you are ready to continue, click the **Proceed** button.

## About your IBMid Account

By creating, or using, an IBMid to access an IBM Web site or other service, you acknowledge that you have read, understand, and agree to be bound by the below terms, in addition to the terms in the IBM Terms of Use, the IBM Privacy Agreement and the “About your IBMid Account Privacy” section of this notice provide details on how IBM protects your personal data.

Most IBMid accounts are individual IDs, meaning you manage your own account information and access your account through the IBMid login service (i.e. IBMid is the authenticating agency).

If your account is created using an email address containing a domain owned by an organization which you are employed by, contracted to, or volunteer for, your organization can:

- inquire about the status of your ID
- request your account settings (including your personal information)
- at its option, convert it to an enterprise ID.

If your account was created as (or is converted to) an enterprise ID, your organization manages your account information. It is also the authenticating agency for your ID, meaning your access is controlled by your organization’s login service.

**Cancel** **Proceed**

At this point, you may see a message on the screen that looks something like this: **Please wait while we create your IBMid and set up the service subscription.**

\_\_\_\_ 6. Once your IBMid and service subscription has been created, you should see a screen that looks something like this:

The screenshot shows a web browser window for the URL [myibm.ibm.com](https://myibm.ibm.com). The page header includes the IBM logo and navigation links for "My IBM", "Profile", and "Billing". A user profile is shown with the email "IBMid: roger.e.sanders@gmail.com". The main content area features a large heading "Roger, your IBM product space is ready!". Below it are links for "Need help?", "Get product support", and a call-to-action button "View catalog". A sidebar at the bottom left contains links for "Contact IBM", "Privacy", "Terms of use", "Accessibility", and "Cookie Preferences". A language selector dropdown shows "United States - English". A blue button labeled "Let's talk" with a speech bubble icon is located in the bottom right corner.

You should also receive an email from *IBM Security* saying your IBMid is activated.

**Congratulations!** You now have an **IBMid**, which you will now use to create an IBM Cloud account (and later, to create an IBM Cloud Pak for Data as a Service account).

## 2. Create an IBM Cloud account.

\_\_\_\_ 1. Open a web browser and go to the *IBM Cloud* web site (<https://www.ibm.com/cloud>). Once there, select the **Create IBM Cloud account** button located in the top right corner of the screen.

The screenshot shows the IBM Cloud homepage. At the top right, there is a blue button labeled "Create IBM Cloud account". A mouse cursor is hovering over this button. The page features a large graphic on the right side with abstract blue and white shapes, including circles and a hexagon. Below the graphic, there are three main sections: "Built for your industry" (with a "Why IBM Cloud" link), "Optimized risk controls" (with a gear icon), and "Workloads anywhere" (with a cloud icon). Each section has a brief description and a "Let's talk" button at the bottom right.

IBM Cloud.  
Hybrid. Open.  
Resilient.

Your platform and partner for digital transformation

Get started for free → Chat with an IBM expert →

Built for your industry  
Why IBM Cloud →

Optimized risk controls

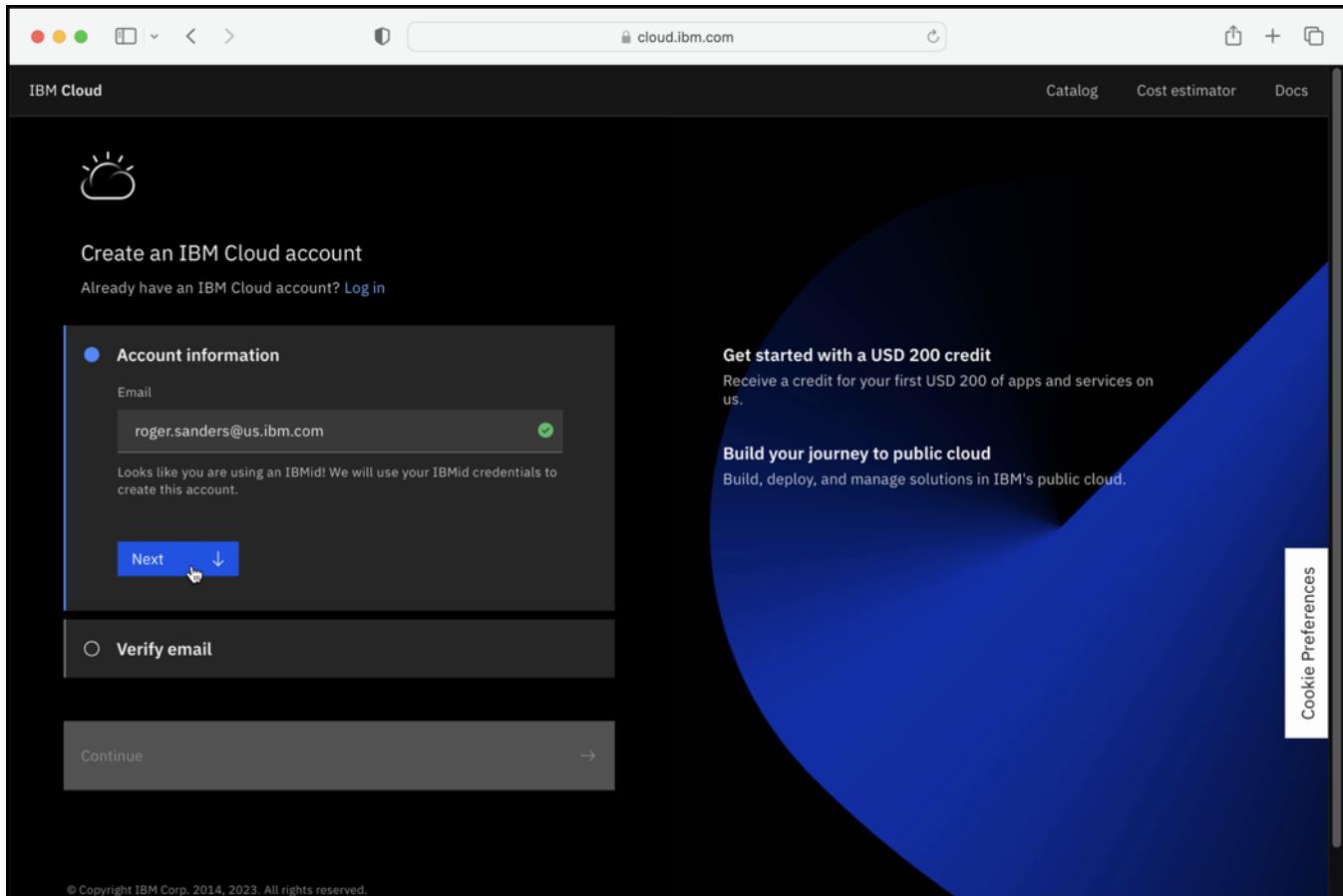
Workloads anywhere

Deploy pre-configured, customized [security and compliance](#) controls across

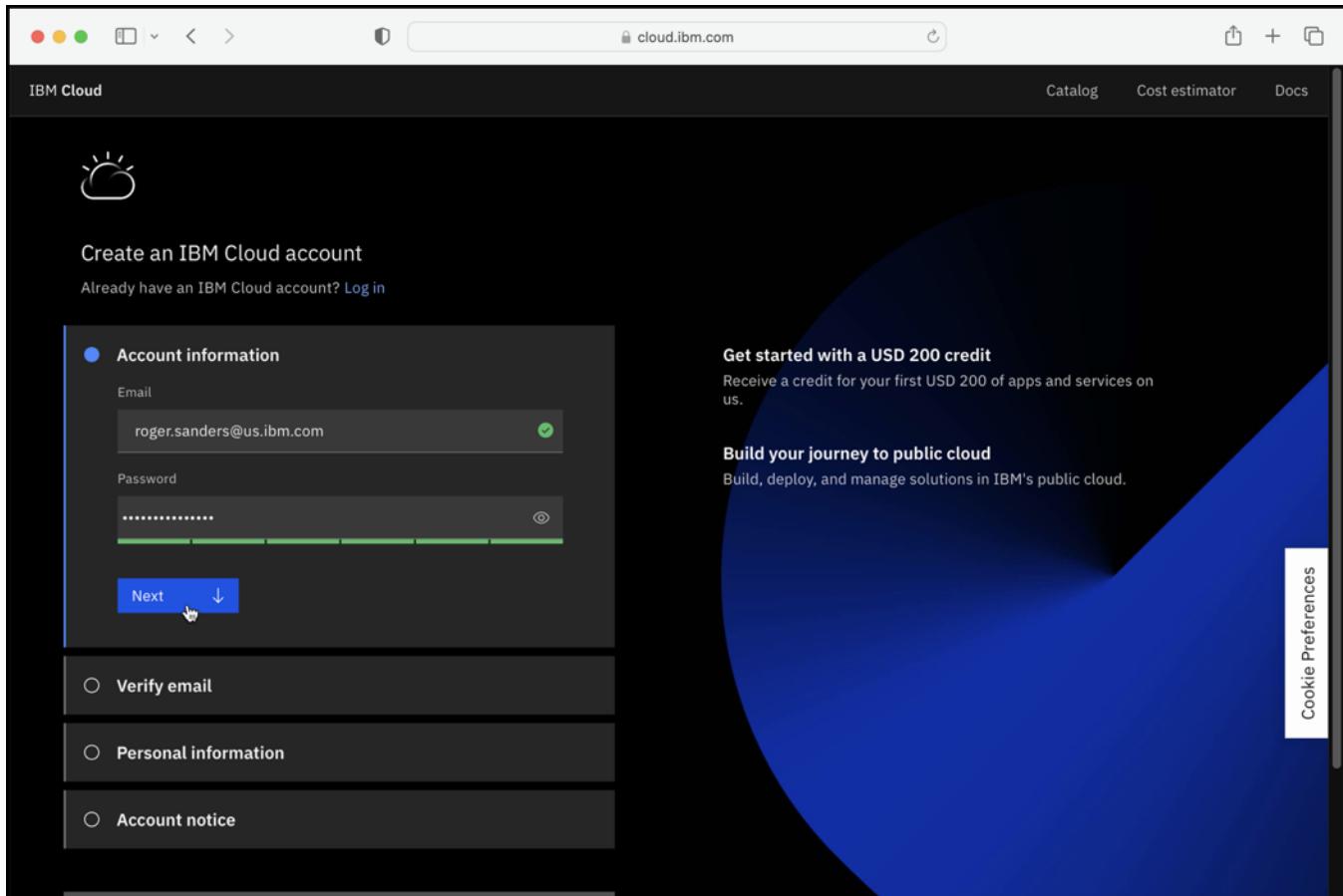
Operate with speed and agility, simplifying how [workloads are distributed](#) and managed, and

Let's talk

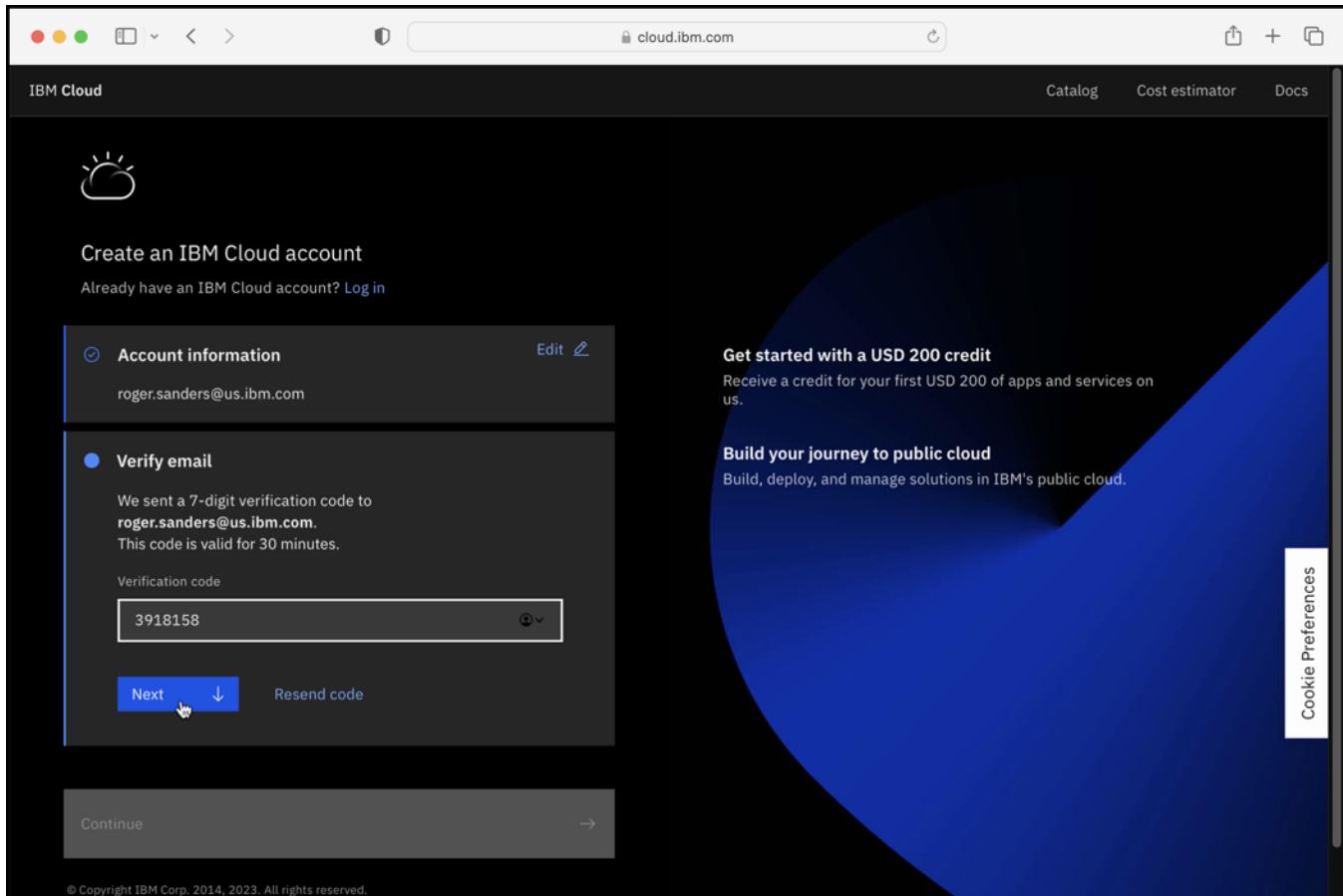
\_\_\_\_ 2. When the *Create an IBM Cloud account* screen is displayed, enter the email address you used to create your IBMid in the **Email** field (which is located inside the **Account information** box). Then, click the **Next** button.



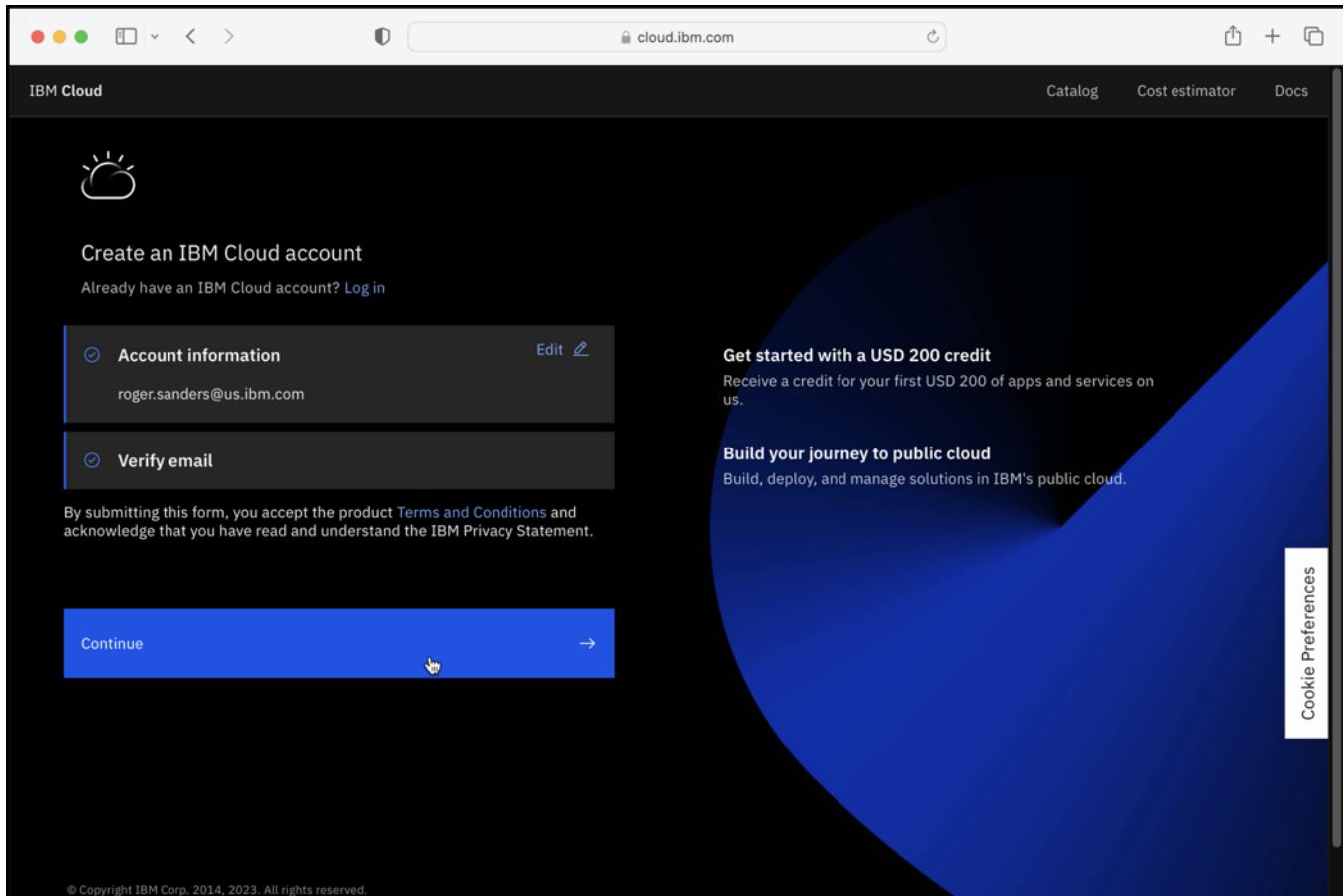
\_\_\_\_ 3. When the **Password** field appears, enter the password associated with the email address (IBMid) provided in the previous step and click the **Next** button. This will cause a 7-digit verification code to be sent to the email address specified. (The email will come from *IBM Cloud* and the subject will be *Your IBM Cloud verification code*.)



\_\_\_\_ 4. Once you receive an email containing a verification code, enter the code in the **Verification code** field and click the **Next** button. (*Note: The verification code is only valid for 30 minutes. If, after 5 minutes, you don't see an email containing the code in your inbox, check your spam/junk email folder.*)



\_\_\_\_ 5. After the verification code you entered in the **Verification code** field is authenticated, the **Continue** button located at the bottom of the screen will be enabled. (Until then, this button is disabled.) When the **Continue** button is enabled, click it. This will cause the current screen to be replaced with the *Review your account privacy notice* screen.

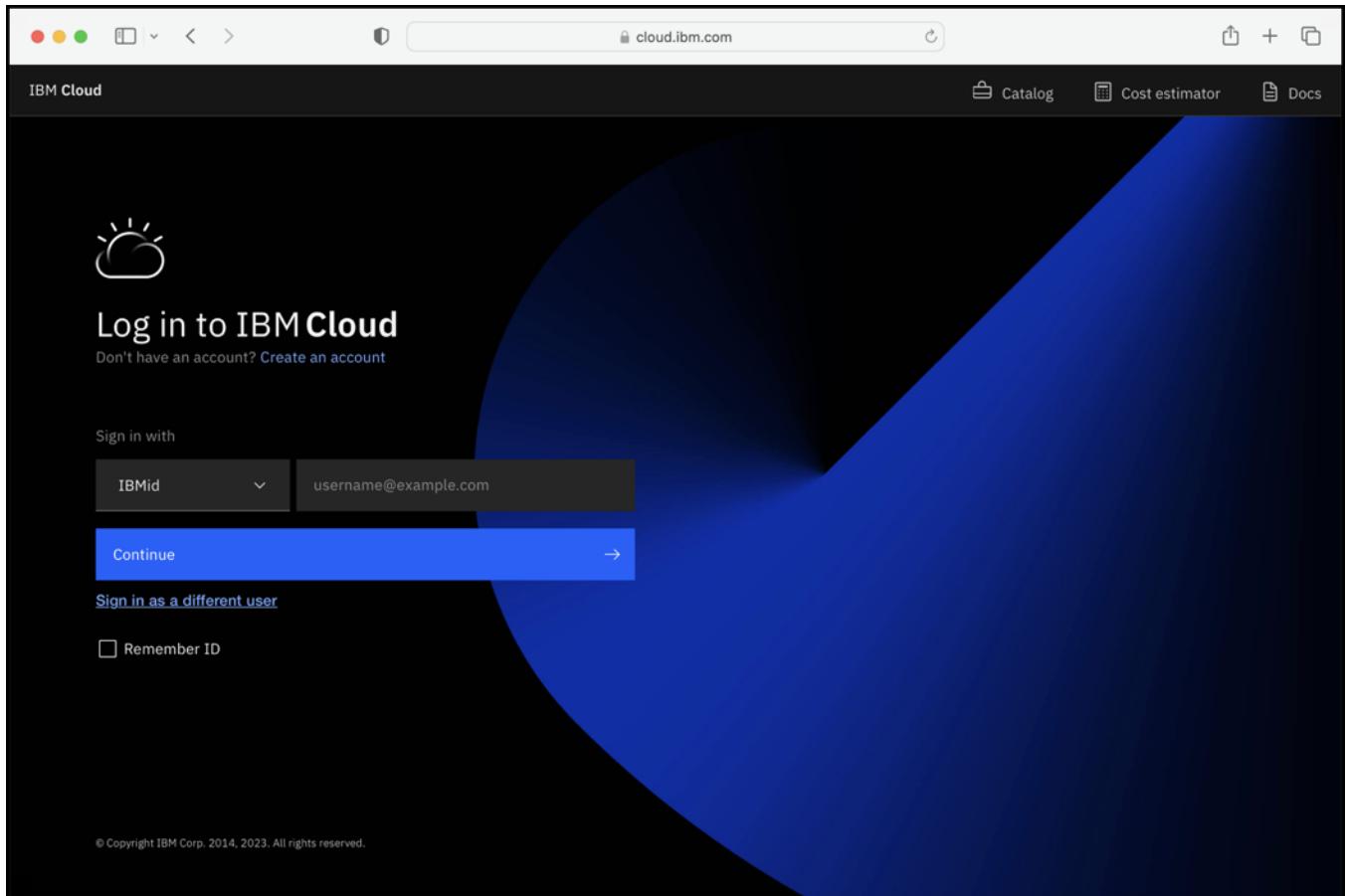


\_\_\_\_ 6. When the *Review your account privacy notice* screen is displayed, read through the terms and conditions of your IBMid Account and then click the **I acknowledge that I understand how IBM is using my Basic Personal Data ...** check-box. Click the **Continue** button to continue.

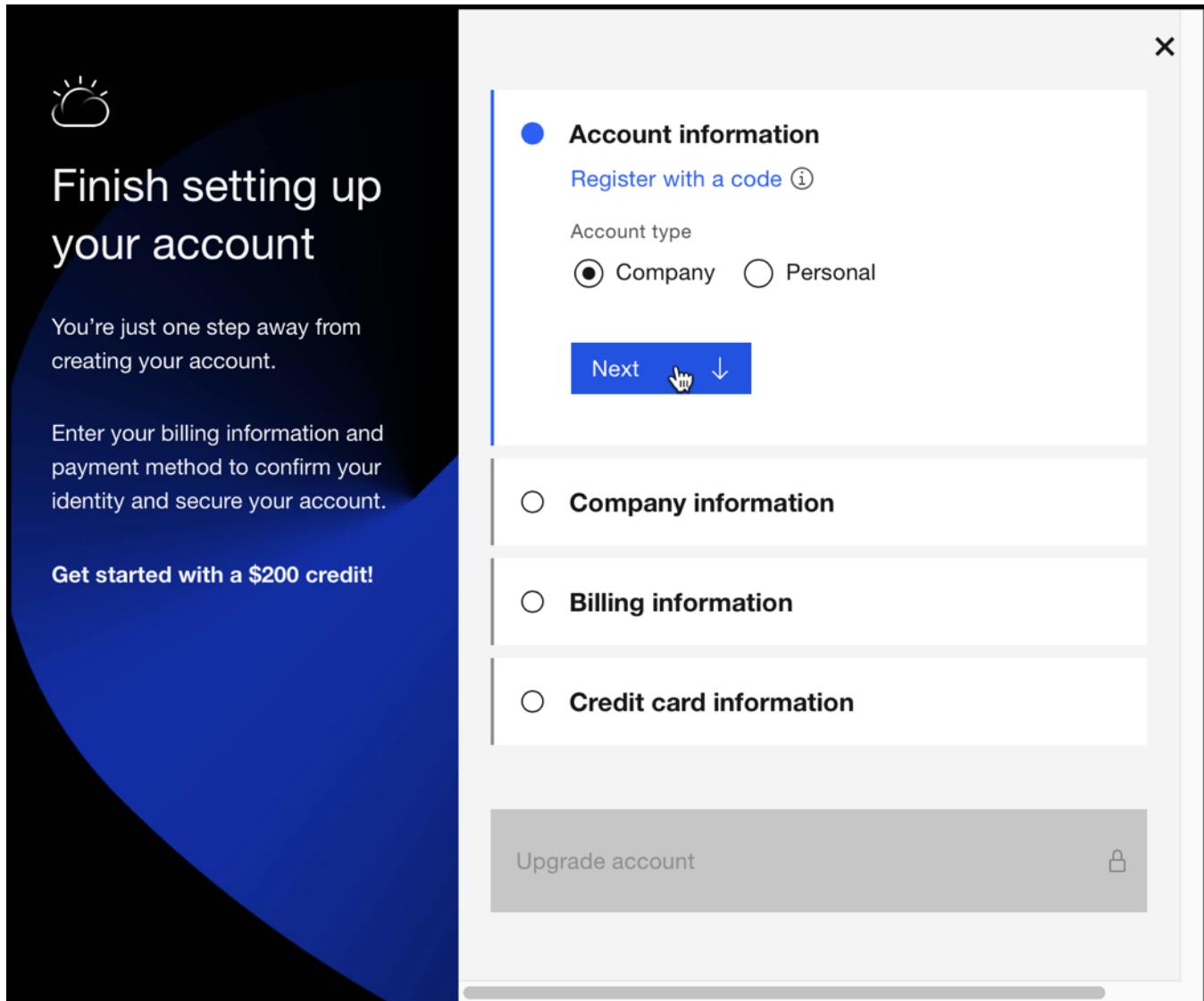
The screenshot shows a web browser window for [cloud.ibm.com](https://cloud.ibm.com). The page title is "IBM Cloud". In the top right, there are links for "Catalog", "Cost estimator", and "Docs". A small "VV" icon is in the top center. The main content area has a dark blue background with a large white circular graphic on the right. At the top left is a small computer monitor icon. Below it, the text "Review your account privacy notice" is displayed. Underneath, a section titled "About your IBMid Account" contains text about account creation and usage. It states that by creating or using an IBMid, you acknowledge reading, understanding, and agreeing to be bound by the terms, including the IBM Privacy Agreement and the "About your IBMid Account Privacy" section. It also notes that most IBMids are individual IDs managed by the user. If created using an email from an organization, the organization manages the account and is the authenticating agency. A list of actions an organization can take is provided. Below this, a note says if your account was converted to an enterprise ID, your organization manages it. A "Last updated: 2021-10-18" timestamp is shown. On the left, a sidebar lists sections with dropdown arrows: "What data does IBM collect?", "Why IBM needs your data", "How your data is obtained", "How IBM uses your data", "How IBM protects your data", "How long we keep your data", and "About your IBMid Account Privacy". At the bottom, a checkbox is checked, stating: "I acknowledge that I understand how IBM is using my Basic Personal Data and (if applicable) how my organization may become the authenticating agency for my IBMid account. I certify that my age is at least the age of consent for my country of residence." A blue "Continue" button with a right-pointing arrow is at the bottom. A "Cookie Preferences" link is on the right side.

The **Continue** button should be replaced with the message “*Creating your account ...*”

7. Once your IBM Cloud account has been created, you will be presented with a *Log in to the IBM Cloud* screen that looks something like this:



\_\_\_\_ 8. Log in to your IBM Cloud account using the IBMid and password you provided earlier. If this is your first time logging into your account, a popup window will appear and you will be asked to provide additional information that's needed to activate the account. Start by clicking the **Company** or **Personal** radio-button to specify the **Account type**. After you have made your selection, click the **Next** button.



\_\_\_\_ 9. If you indicated that your IBM Cloud account should be treated as a **Company** account, provide the appropriate **Company information** (i.e., **Company name** and company contact information) in the fields shown.



# Finish setting up your account

You're just one step away from creating your account.

Enter your billing information and payment method to confirm your identity and secure your account.

**Get started with a \$200 credit!**

Company information

Company name: IBM

Country or region: United States

Address line 1: 2100 Stewart Street

Address line 2 (optional): The apartment, suite, unit, building or floor

City: Seattle

State: North Carolina

Zip code: 27026-0000

Phone number: +1 800 720-0776

X

\_\_\_\_ 10. After you have provided all of the **Company information** requested (assuming this is a Company account), click the **Next** button.



## Finish setting up your account

You're just one step away from creating your account.

Enter your billing information and payment method to confirm your identity and secure your account.

Get started with a \$200 credit!

City: Puppy Venn

State: North Carolina Zip code: 27028-8888

Phone number: +1 800 729-3776

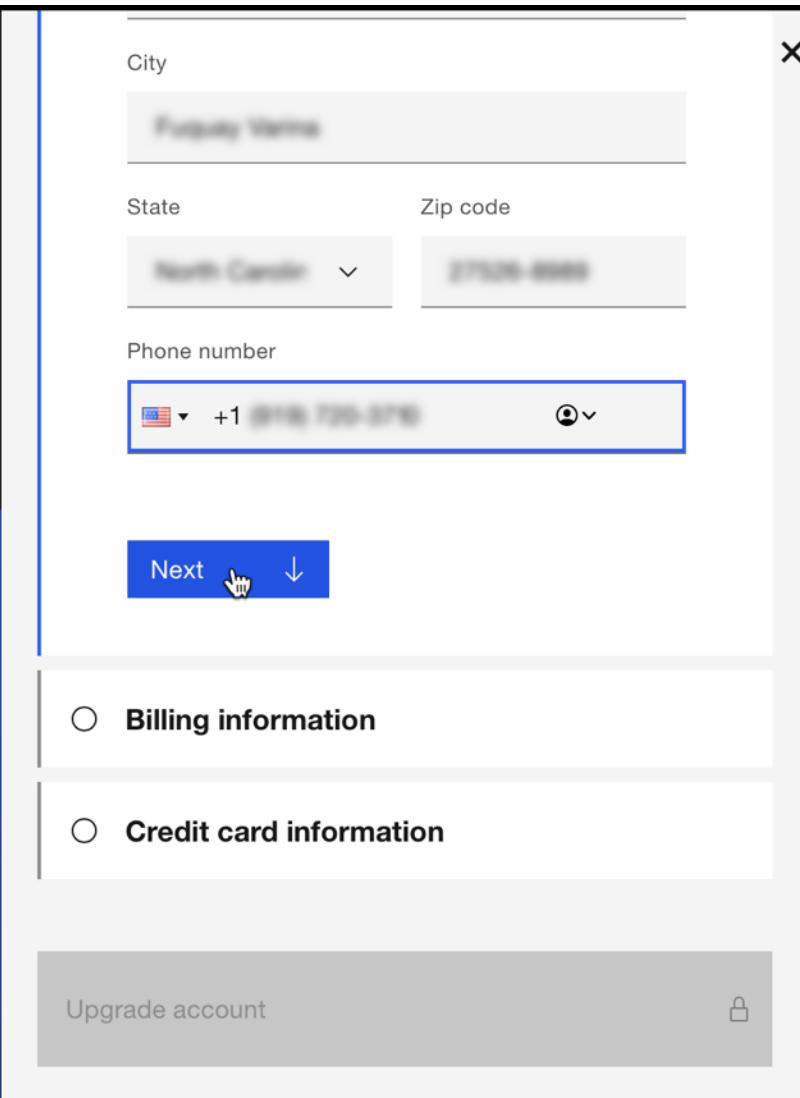
Next 

**Billing information**

**Credit card information**

Upgrade account 

**IBM**



\_\_\_\_ 11. Now provide the appropriate **Billing information**. If this is a **Company** account and you are the billing contact, click the **My billing contact is the same as my company contact** check-box. Otherwise, provide the appropriate **Billing information** (i.e., contact information for the individual responsible for paying any charges incurred) in the fields shown.



## Finish setting up your account

You're just one step away from creating your account.

Enter your billing information and payment method to confirm your identity and secure your account.

**Get started with a \$200 credit!**

**Billing information**

First name  
Roger

Last name  
Sanders

My billing contact is the same as my company contact

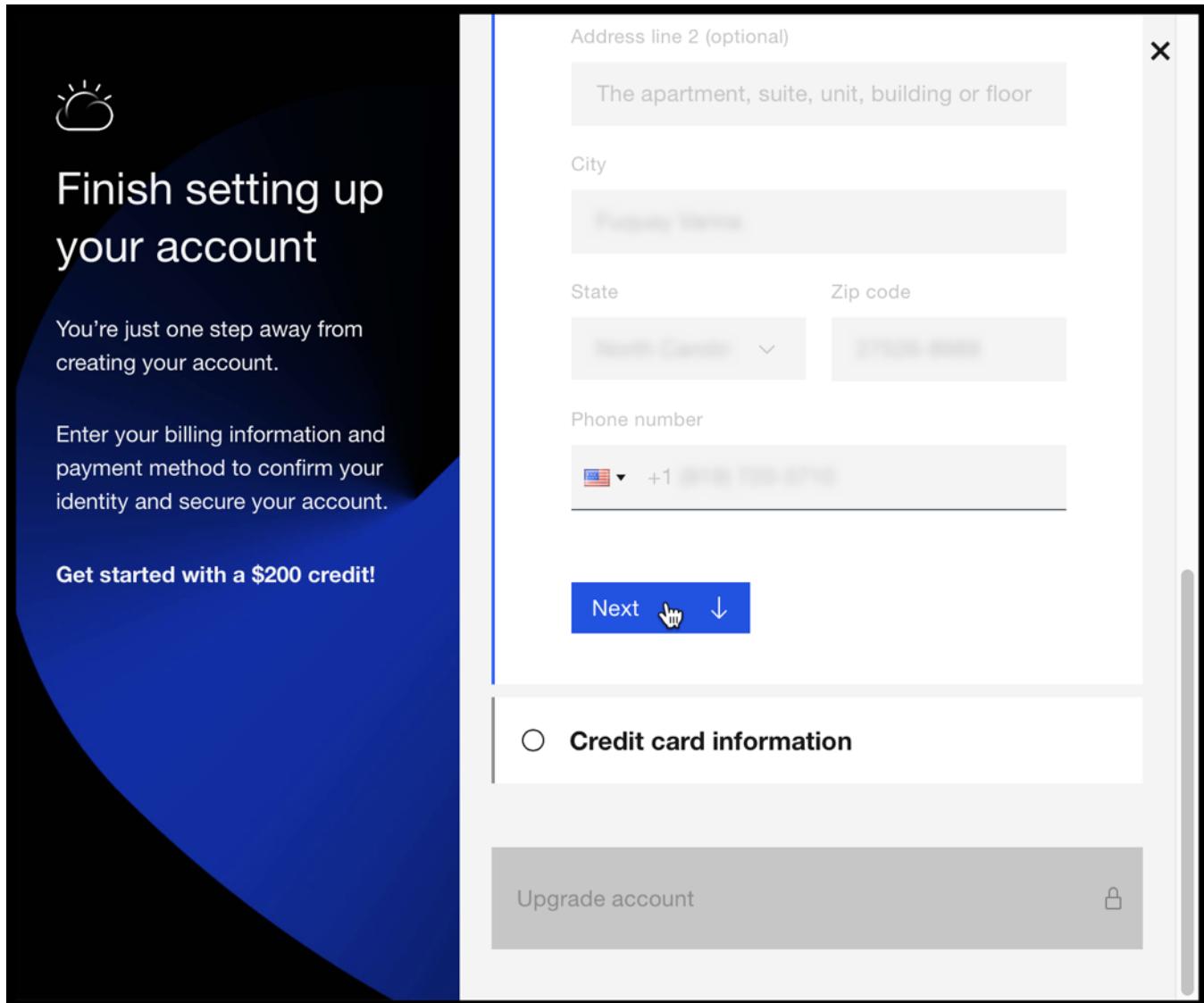
Country or region  
United States

Address line 1  
[REDACTED] Search for new address

Address line 2 (optional)  
The apartment, suite, unit, building or floor

City  
[REDACTED]

\_\_\_\_ 12. Once you have confirmed the appropriate **Billing information** has been provided, click the **Next** button.



The screenshot shows the 'Finish setting up your account' step of an IBM Cloud account creation process. On the left, a dark blue sidebar displays a cloud icon and the text: 'Finish setting up your account', 'You're just one step away from creating your account.', 'Enter your billing information and payment method to confirm your identity and secure your account.', and 'Get started with a \$200 credit!'. The main right area contains fields for address line 2 (optional), apartment/suite information, city, state, zip code, and phone number. A 'Next' button with a downward arrow is at the bottom. Below it, a section titled 'Credit card information' is shown with an unchecked radio button. At the very bottom is an 'Upgrade account' button with a lock icon.

Address line 2 (optional)

The apartment, suite, unit, building or floor

City

State Zip code

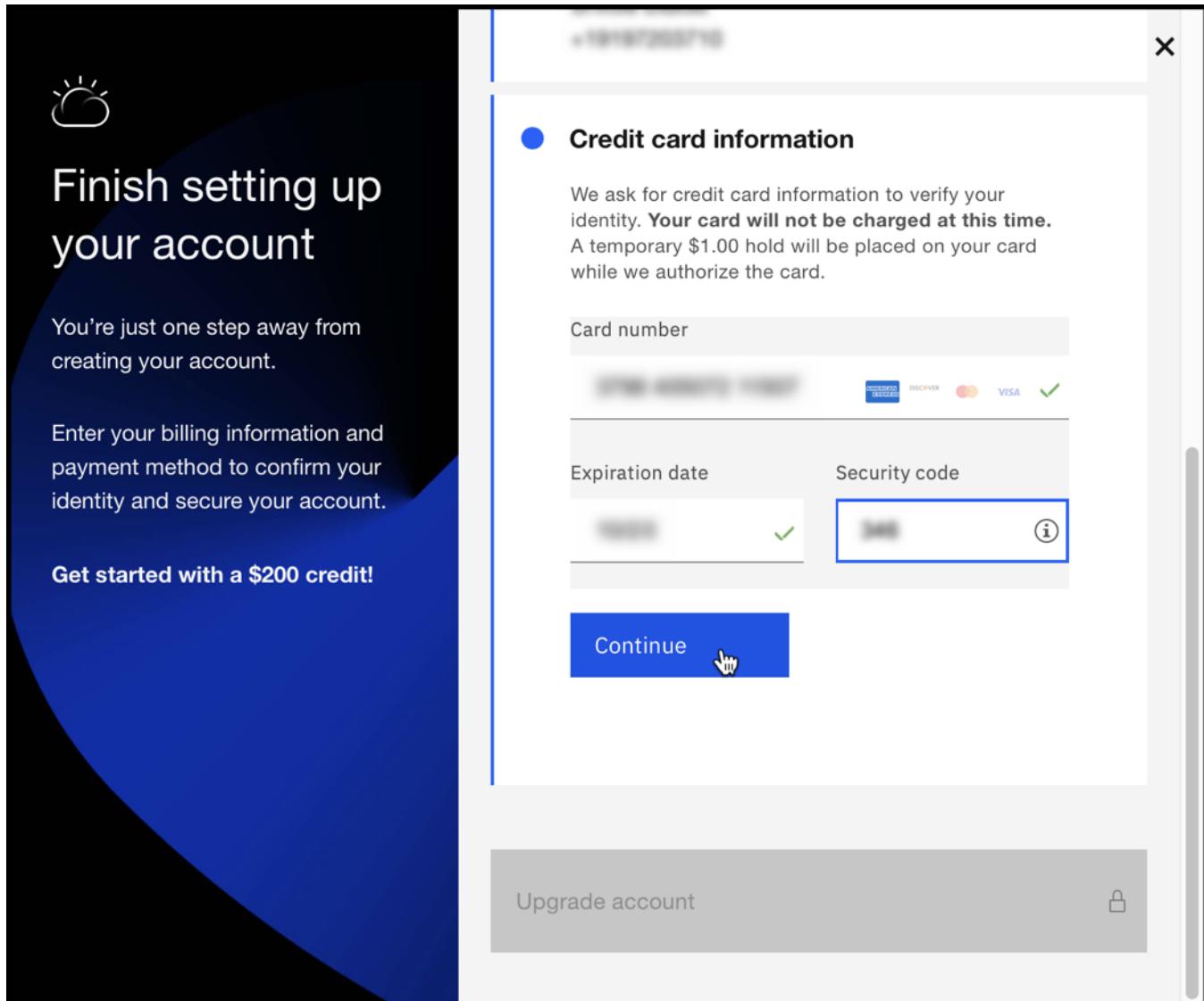
Phone number

Next  ↓

Credit card information

Upgrade account 

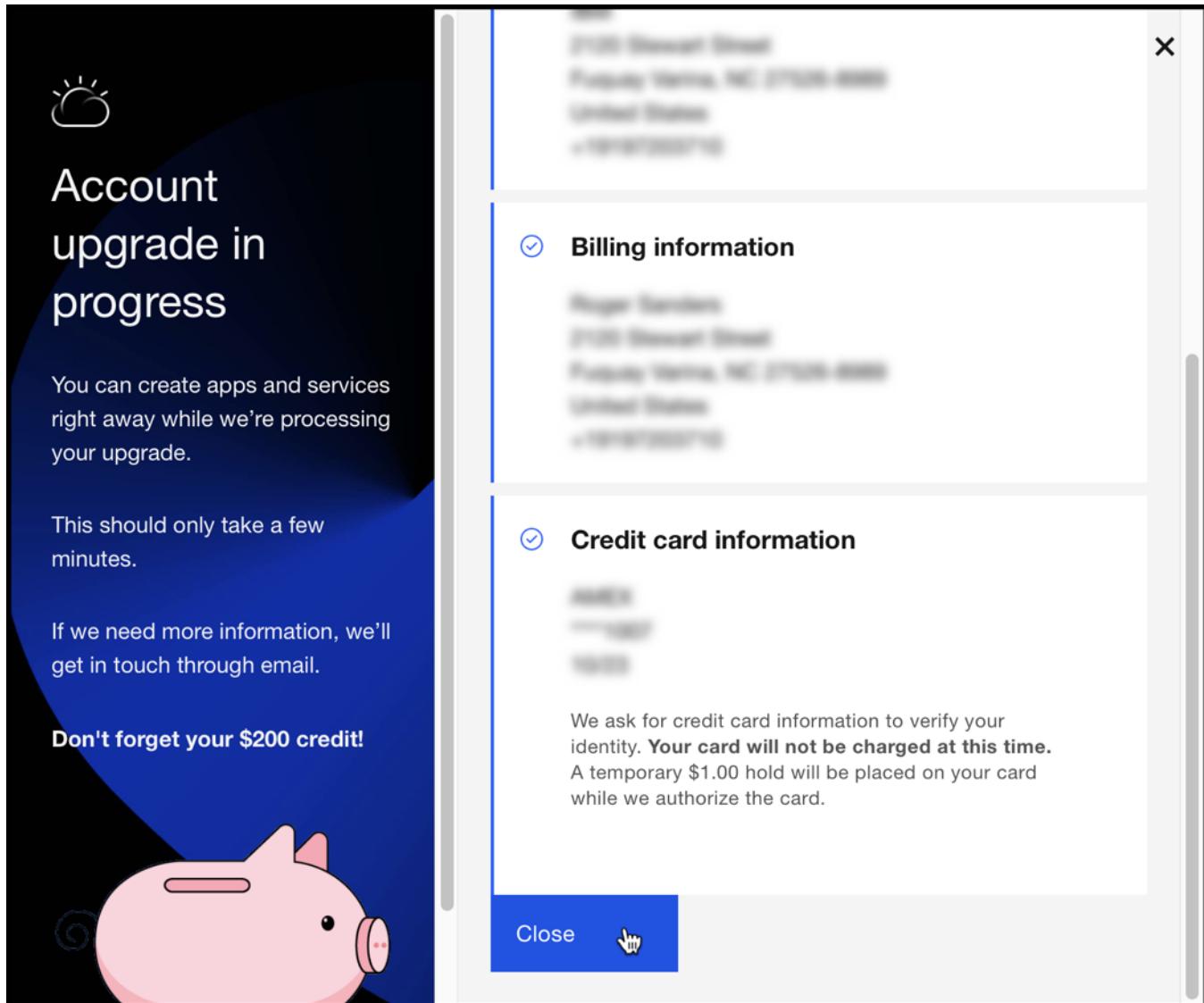
13. Now, enter the appropriate **Credit card information** (i.e., **Card number**, **Expiration date**, and **Security code**) in the fields provided. It is important to note that *this information is only used to verify your identity (therefore, the name on the credit card used MUST match the name provided for the billing contact)*. Your card will **NOT** be charged unless you decide to purchase a billable IBM Cloud service. **Because this workshop uses the no-charge IBM Db2 "Lite" service and the IBM Watson Studio "Lite" service, NO charges will be incurred.** After you have provided the credit card information requested, click the **Continue** button. This will cause the **Upgrade account** button located at the bottom of the screen to be enabled.



\_\_\_\_ 14. Once the **Upgrade account** button is enabled, click it. This will cause a temporary \$1.00 US hold to be placed on the credit card provided for authorization purposes. (This will also cause the **Upgrade account** button to be replaced with a message saying, “**Authorizing your credit card ...**”.)

The screenshot shows a step in the account setup process. On the left, a dark sidebar features a sun icon and the text "Finish setting up your account". Below it, a message says "You're just one step away from creating your account." and "Enter your billing information and payment method to confirm your identity and secure your account." A blue button at the bottom left says "Get started with a \$200 credit!". The main area has a white background with a blue header bar. The header bar contains the name "Roger Sanders" and a close button "X". Below the header is a section titled "Credit card information" with an "Edit" link. A note states: "We ask for credit card information to verify your identity. Your card will not be charged at this time. A temporary \$1.00 hold will be placed on your card while we authorize the card." A callout box provides additional information: "Your card will not be charged at this time. A temporary \$1.00 hold will be placed on your AMEX while we authorize the card." At the bottom is a blue button labeled "Upgrade account" with a hand cursor icon.

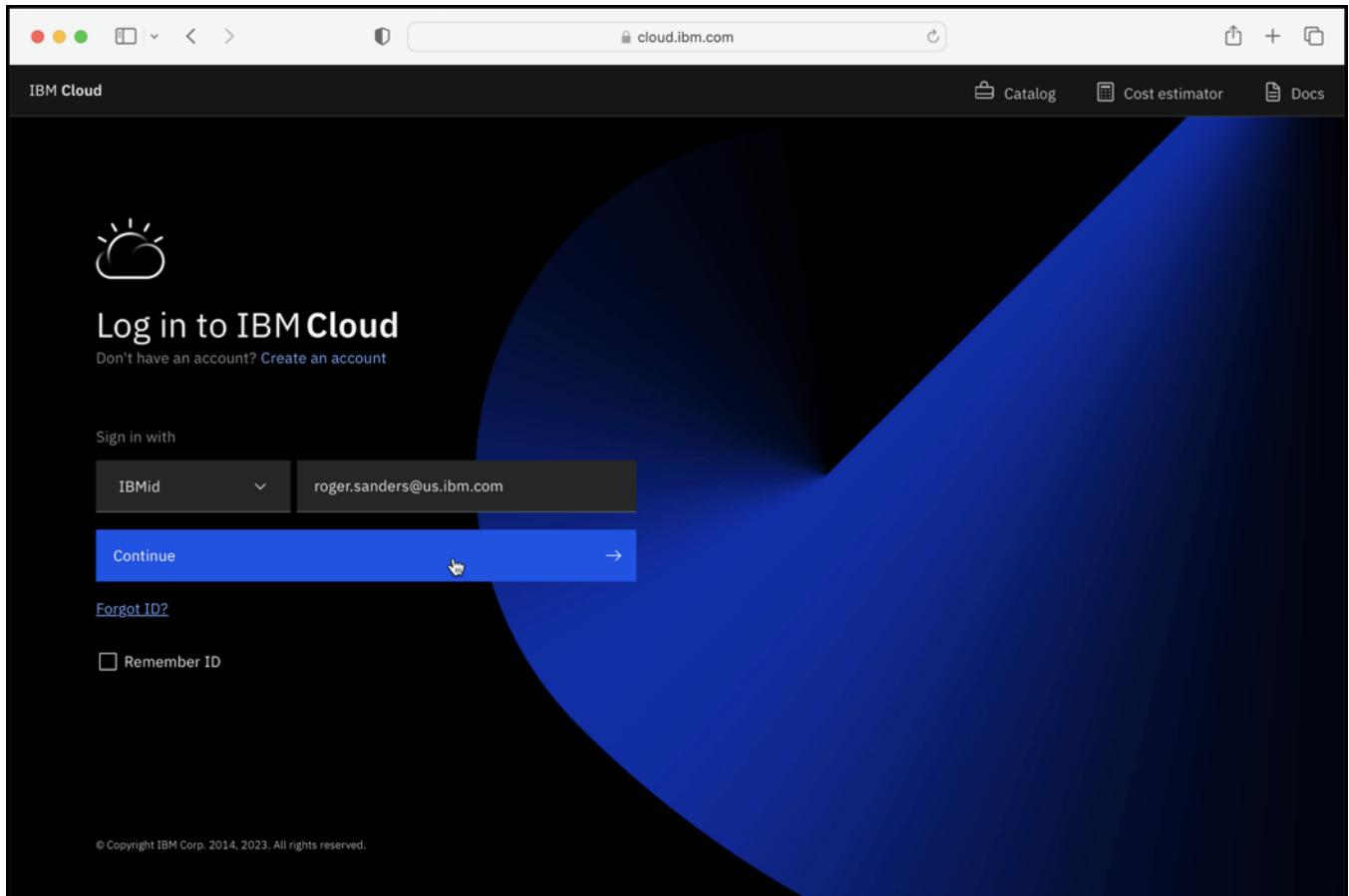
\_\_\_\_ 15. When the credit card authorization process is complete, a **Close** button will be displayed at the bottom of the popup window. Click on this button. (When your IBM Cloud account is ready, you should receive an email from *IBM Cloud* saying “Welcome to IBM Cloud! Let’s get started.”)



**Congratulations!** You now have an active **IBM Cloud** account that can be used to perform the rest of the exercises in this workshop.

### 3. Provision a Db2 on Cloud Lite plan database

- \_\_\_\_ 1. Open a web browser and go to the **IBM Cloud** login web site (<https://cloud.ibm.com/login>). Once there, ensure the **Sign in with** dropdown list contains the value **IBMid** and enter the email address associated with your IBMid in the field immediately beside the dropdown list. When finished, click the **Continue** button. (*Note: If you do not have a valid IBMid and an IBM Cloud account, complete steps 1 and 2 of this exercise before continuing.*)



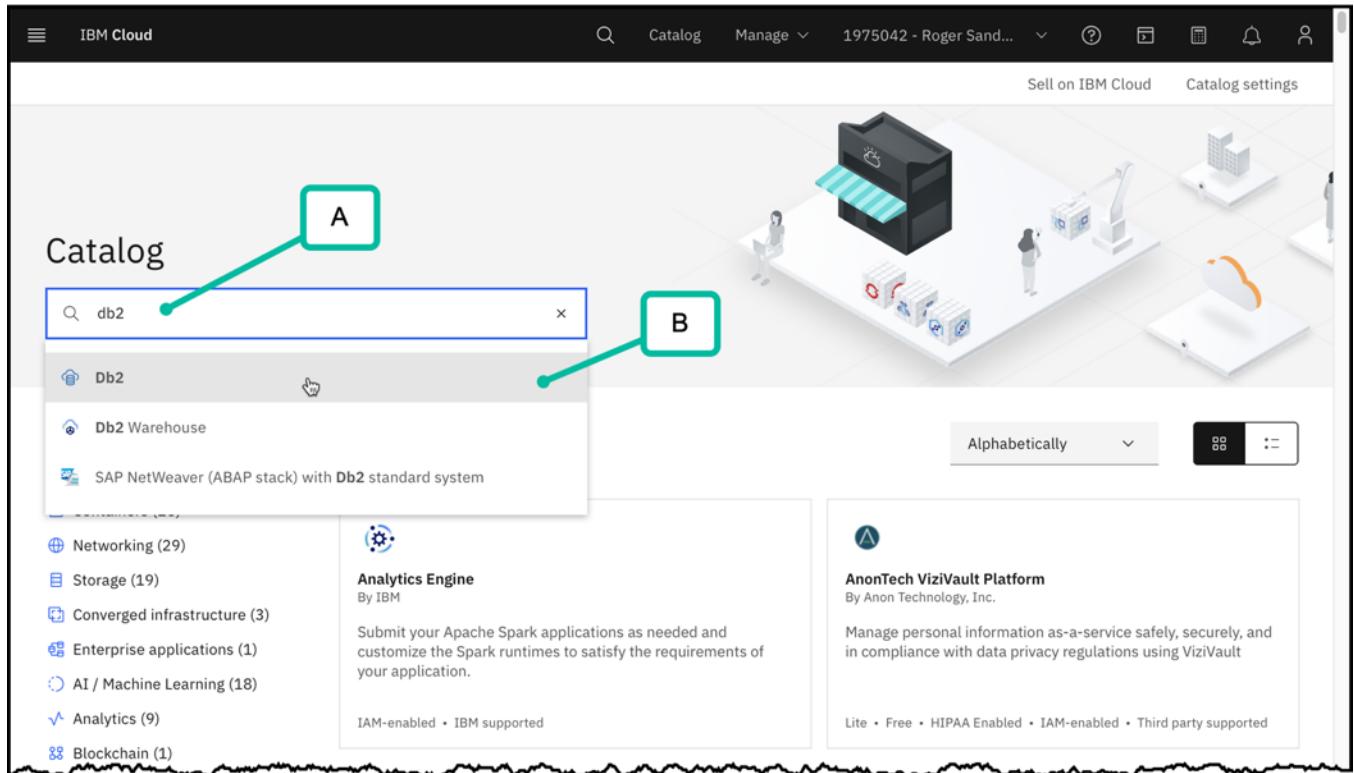
\_\_\_\_ 2. If prompted, enter the password associated with the IBMid provided. Once you are logged in, you should see a screen that looks something like this:

The screenshot shows the IBM Cloud dashboard. At the top right, there is a blue button labeled "Create resource" with a plus sign. To its right is a dropdown menu labeled "Select an option". Below this, there are several cards: "Build" (Explore IBM Cloud with this selection of easy starter tutorials and services), "Create a custom dashboard" (Create a shareable dashboard that you can customize with widgets, scope, and your own layout), "Choose a Database" (Find the right IBM Cloud database for the job), "Databases for MongoDB" (Get MongoDB's rich JSON query and aggregation features with autoscaling and automatic backups with IBM Cloud™ Databases for MongoDB), and "Databases for Elasticsearch" (Get IBM Cloud™ Databases for Elasticsearch, a JSC database with powerful text search, high availability, and backups). On the left, there is a sidebar with various icons for services like Cloud Functions, Cloud Storage, and Cloud Databases.

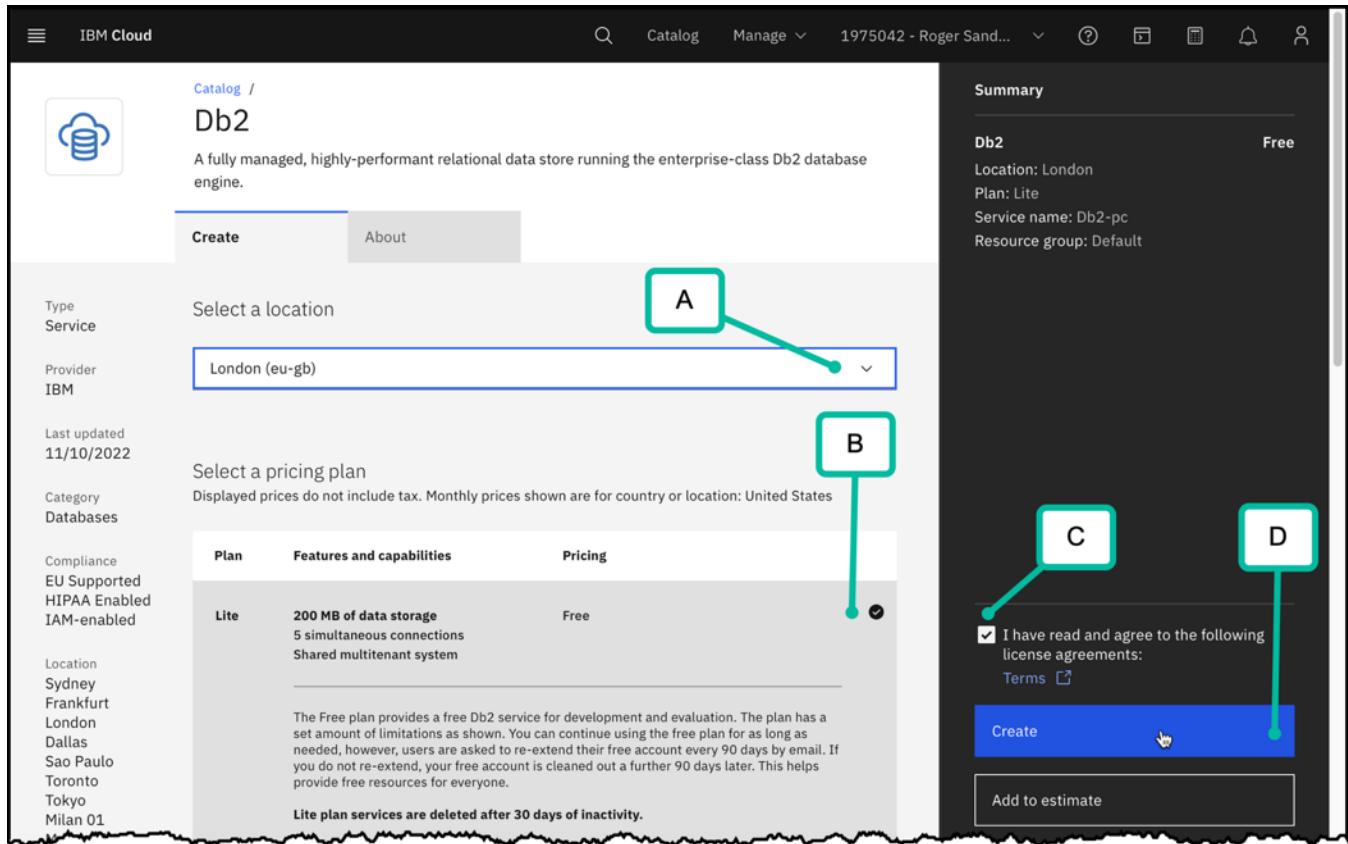
\_\_\_\_\_ 3. Click on the **Create resource** button located at the top right corner of the screen. This will cause the IBM Cloud Catalog screen to be displayed.

This screenshot is identical to the one above, but the "Create resource" button is highlighted with a mouse cursor, indicating it is being clicked. The rest of the interface remains the same, showing the dashboard cards and sidebar.

\_\_\_\_ 4. When the IBM Cloud *Catalog* screen appears, [A] type the characters “db2” in the **Search** field located in the top left corner of the screen (i.e., the entry field with the magnifying glass icon on the left.) This will cause a list of the Db2 resources available on IBM Cloud to be displayed. [B] Click on the item **Db2** in this list. This will cause the *Db2* resource screen to be displayed.



\_\_\_\_ 5. On the *Db2* resource screen, [A] use the **Select a location** dropdown list to choose the location where the Db2 on Cloud database is to be created. For this workshop, either the **Dallas (us-south)** or the **London (eu-gb)** location must be chosen – that’s because these are the only locations that offer the Db2 on Cloud “Lite” (free) plan. [B] Next, choose the **Lite** plan from the **Select a pricing plan** list shown on the screen. [C] Once you have made these selections, indicate that you have read and agree to the license agreements for using Db2 on Cloud by clicking on the **I have read and agree ...** check-box. This will enable the **Create** button. (The license agreements can be reviewed by clicking on the **Terms** link located under this check-box.) [D] When ready, click the **Create** button.



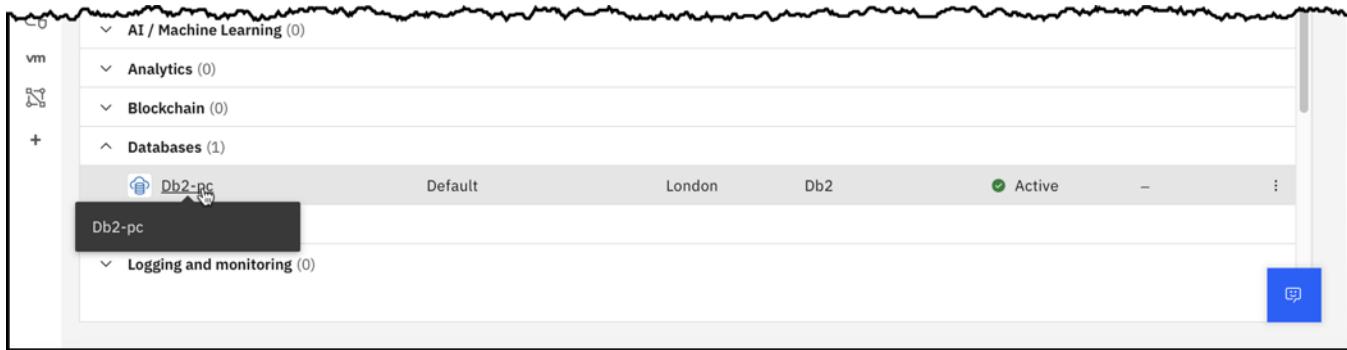
6. When the **Create** button on the *Db2* resource screen is pressed, it will disappear and the message “**Creating ...**” will be displayed in its place. Shortly afterwards, the *Db2* resource screen will be replaced with an IBM Cloud **Resource list** screen. This screen contains a list of every resource/service that has been provisioned under your IBM Cloud account. If you expand the **Databases** item in this list, the *Db2* on Cloud database just provisioned can be seen. Notice that the **Status** column for this database contains the value “**Provision in progress.**”

The screenshot shows the IBM Cloud Resource list interface. On the left, there's a sidebar with various icons for Compute, Containers, Networking, Storage, Converged infrastructure, Enterprise applications, AI / Machine Learning, Analytics, Blockchain, and Databases. The 'Databases' section is expanded, showing one item: 'Db2-pc'. The main table has columns for Name, Group, Location, Product, Status, and Tags. The 'Db2-pc' row is selected, showing details: Group 'Default', Location 'London', Product 'Db2', Status 'Provision in progress', and a 'More' button. There are also filter bars at the top for Name, Group, Location, Product, Status, and Tags.

\_\_\_\_ 7. Refresh the web browser screen periodically until the **Status** column for the new Db2 on Cloud database contains the value “**Active**.” At this point, your database is ready for use.

This screenshot shows the same IBM Cloud Resource list interface as the previous one, but the database 'Db2-pc' now has a green 'Active' status indicator in the Status column. The rest of the table and sidebar are identical to the first screenshot.

\_\_\_\_ 8. Access the newly created Db2 on Cloud database by clicking on the database ID shown in the **Resource list**. (In this example, the database ID is *Db2-pc*. As we will see later, the name of the database is *bludb*.)



\_\_\_\_\_ 9. Clicking on a resource ID in the **Resource list** will cause the screen to be replaced with a high-level “management” screen for the resource selected. (In this example, the resource selected is the Db2 on Cloud database with the ID *Db2-pc*). If the resource chosen was a Db2 on Cloud database, once this “management” screen appears, you can begin working directly with the database, [but only through the Db2 on Cloud User Interface \(UI\)](#).

IBM Cloud

Resource list / Db2-pc Active Add tags

Details Actions...

Manage	Getting started	Need help?
Getting started Service credentials Connections	<p>Where can I find my credentials? Get your username and password by clicking the "Service Credentials" link to the left and selecting "New Credentials". Don't see this menu on the left? Click on "Manage in IBM Cloud" to open the IBM Cloud dashboard.</p> <p><a href="#">Go to UI</a> <a href="#">Getting started docs</a></p>	<p>Submit a IBM Cloud Support Case to our team.</p> <p><a href="#">Support case</a></p>

## 4. Create a service credential for the Db2 on Cloud database.

\_\_\_\_\_ 1. To work with an IBM Cloud resource from another application (for example, a Jupyter Notebook), one or more service credentials for that resource must exist. As this screen illustrates, when a Db2 on Cloud database is first provisioned, it has **No service credentials**. To create a service credential for the database, [A] click on the

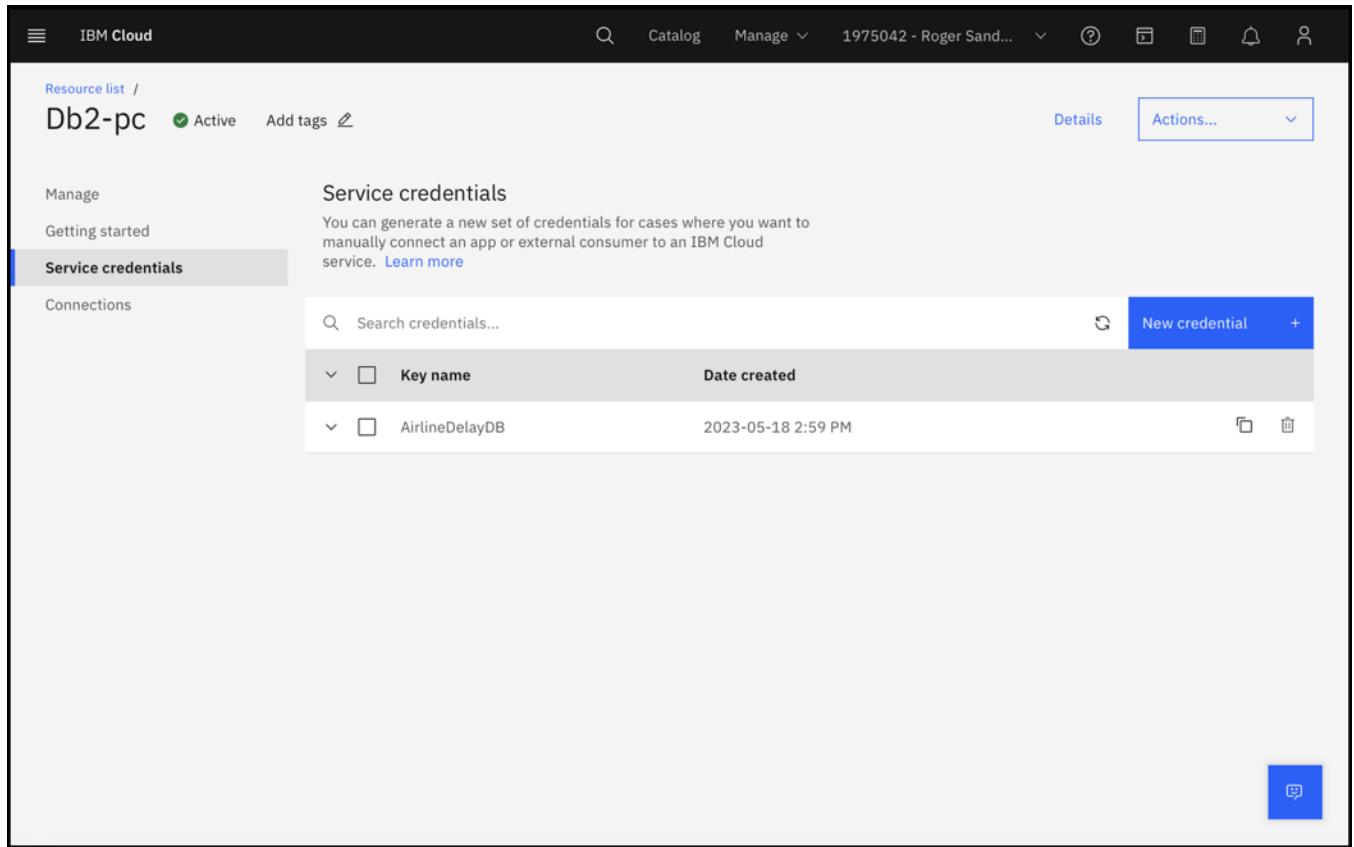
**Service credentials** item in the menu shown in the top left corner of the screen. [B] Then, click the **New credential** button located on the right. This will cause the *Create credential* popup window to be displayed.

The screenshot shows the IBM Cloud Service credentials page for a resource named 'Db2-pc'. The 'Service credentials' tab is selected, indicated by a green box labeled 'A'. On the right, there is a 'New credential' button with a '+' icon, also highlighted with a green box labeled 'B'. A green arrow points from 'B' to the button.

2. When the *Create credential* popup window is displayed, [A] enter a name for the new service credential in the **Name:** field. (The default name will be *Service credentials-1*, but for this exercise, we're going to change the name to *AirlineDelayDB*). [B] Then, click the **Add** button located at the bottom right corner of the window. (Keep the **Role:** of *Manager*.)

The screenshot shows the 'Create credential' dialog box. The 'Name:' field contains 'AirlineDelayDB', which is highlighted with a green box labeled 'A'. At the bottom right of the dialog, there is an 'Add' button highlighted with a green box labeled 'B'. A green arrow points from 'B' to the 'Add' button.

\_\_\_\_ 3. When the *Create credential* popup window closes, the new service credential will be shown in the list of service credentials available.



The screenshot shows the IBM Cloud interface for managing resources. The top navigation bar includes 'IBM Cloud', 'Catalog', 'Manage', and a user profile. Below the navigation is a 'Resource list' for 'Db2-pc'. The 'Service credentials' tab is selected, showing a list of credentials. A search bar at the top right says 'Search credentials...'. A blue button labeled 'New credential' is visible. The list contains one item:

Key name	Date created
AirlineDelayDB	2023-05-18 2:59 PM

Each row has a copy icon (blue square with white text) on the right side.

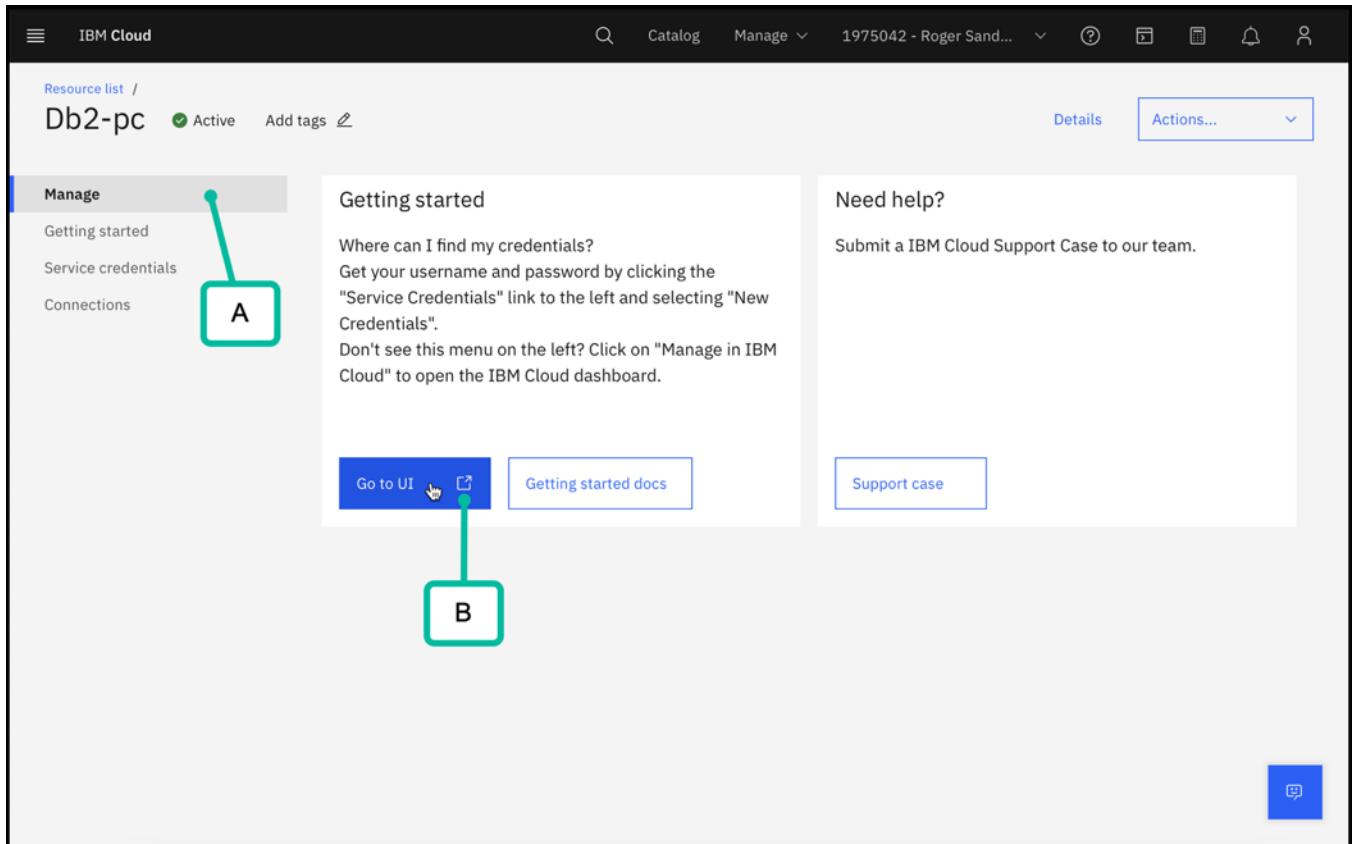
\_\_\_\_ 4. Click the “**Copy to clipboard**” icon shown on the right-hand side of the listing for the service credential just created. Then, paste the information copied into a text file – for this workshop, use a file named **AirlineDelayDB\_Credentials.txt** – and save it for later.

The screenshot shows the IBM Cloud Service credentials page for a resource named 'Db2-pc'. The 'Service credentials' tab is selected. A table lists two entries: 'Key name' (AirlineDelayDB) and 'Date created' (2023-05-18 2:59 PM). A 'Copied!' message is visible next to the delete icon for the second entry.

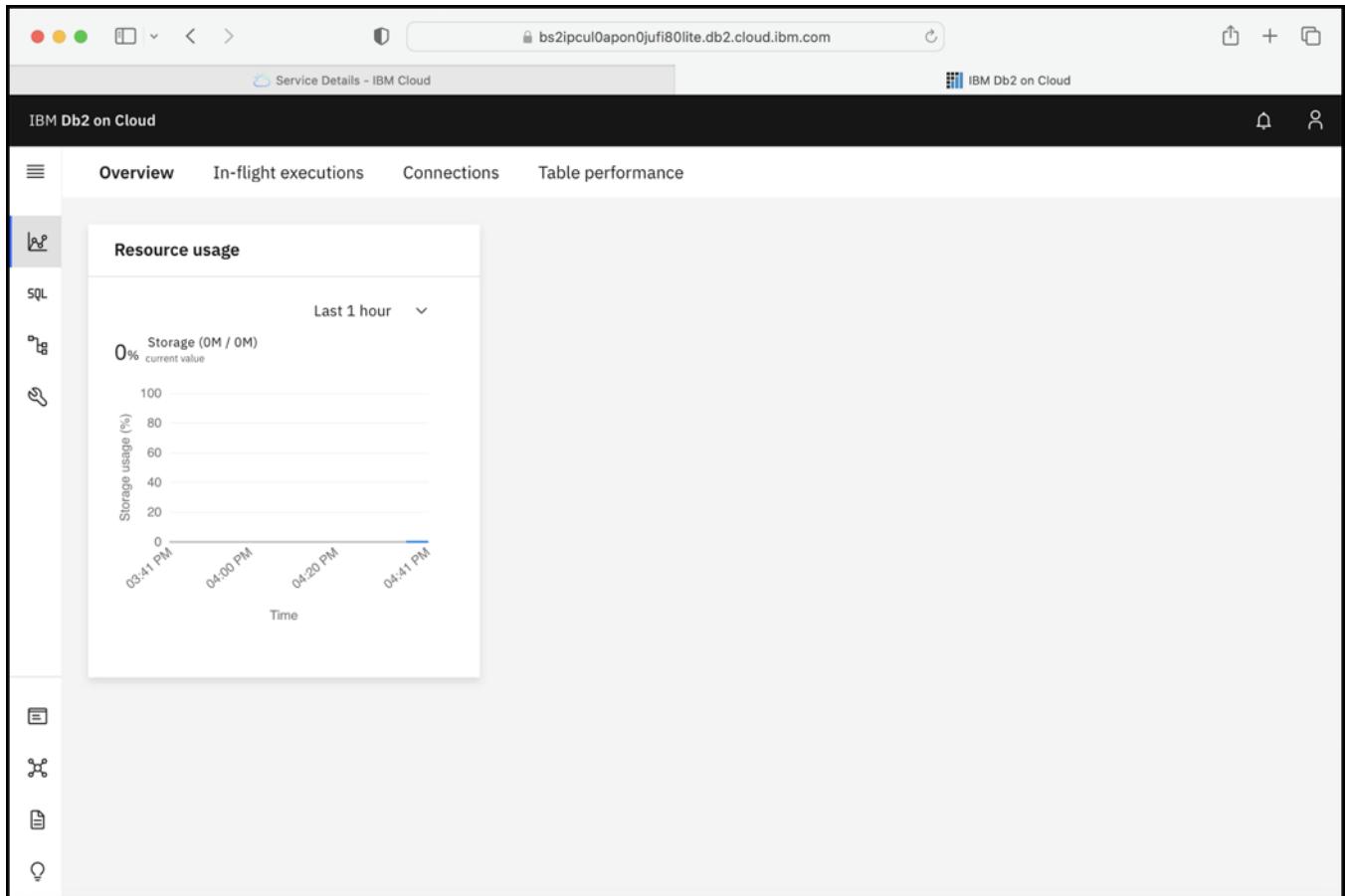
Key name	Date created
AirlineDelayDB	2023-05-18 2:59 PM

## 5. Query the system catalog using the Db2 on Cloud User Interface (UI).

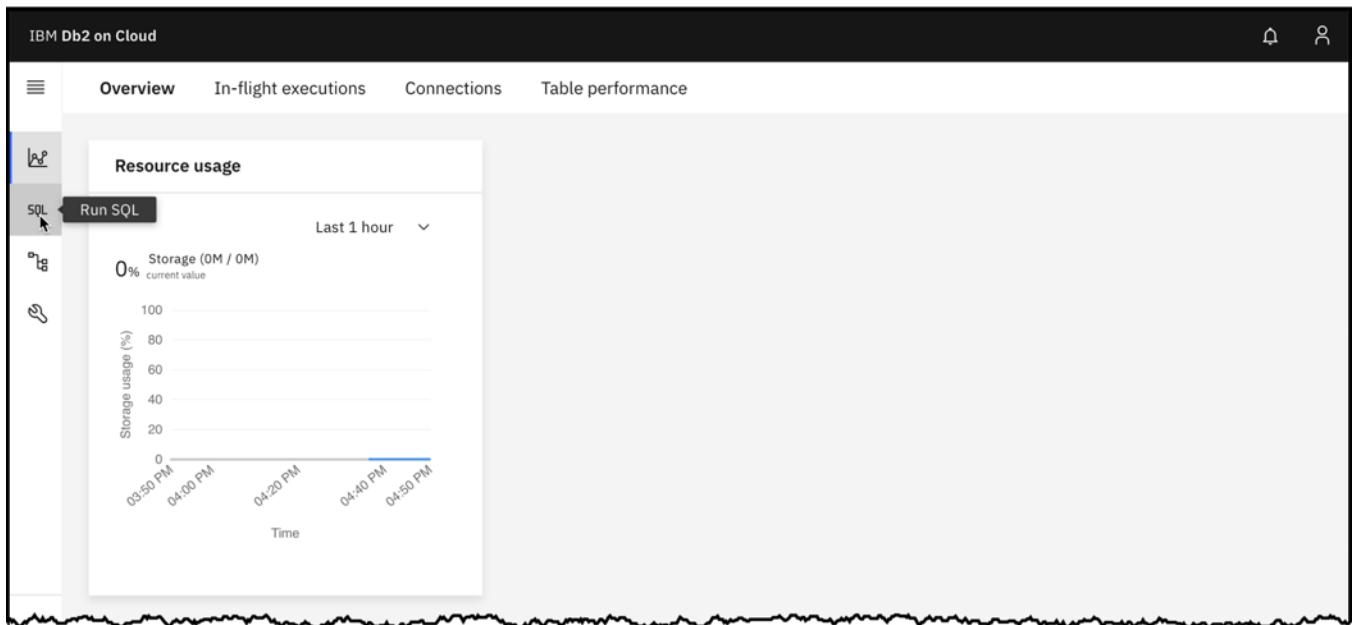
- \_\_\_\_ 1. To work with the Db2 on Cloud database using the Db2 on Cloud User Interface, [A] click on the **Manage** item in the menu shown in the top left corner of the screen. [B] Then, click the **Go to UI** button.



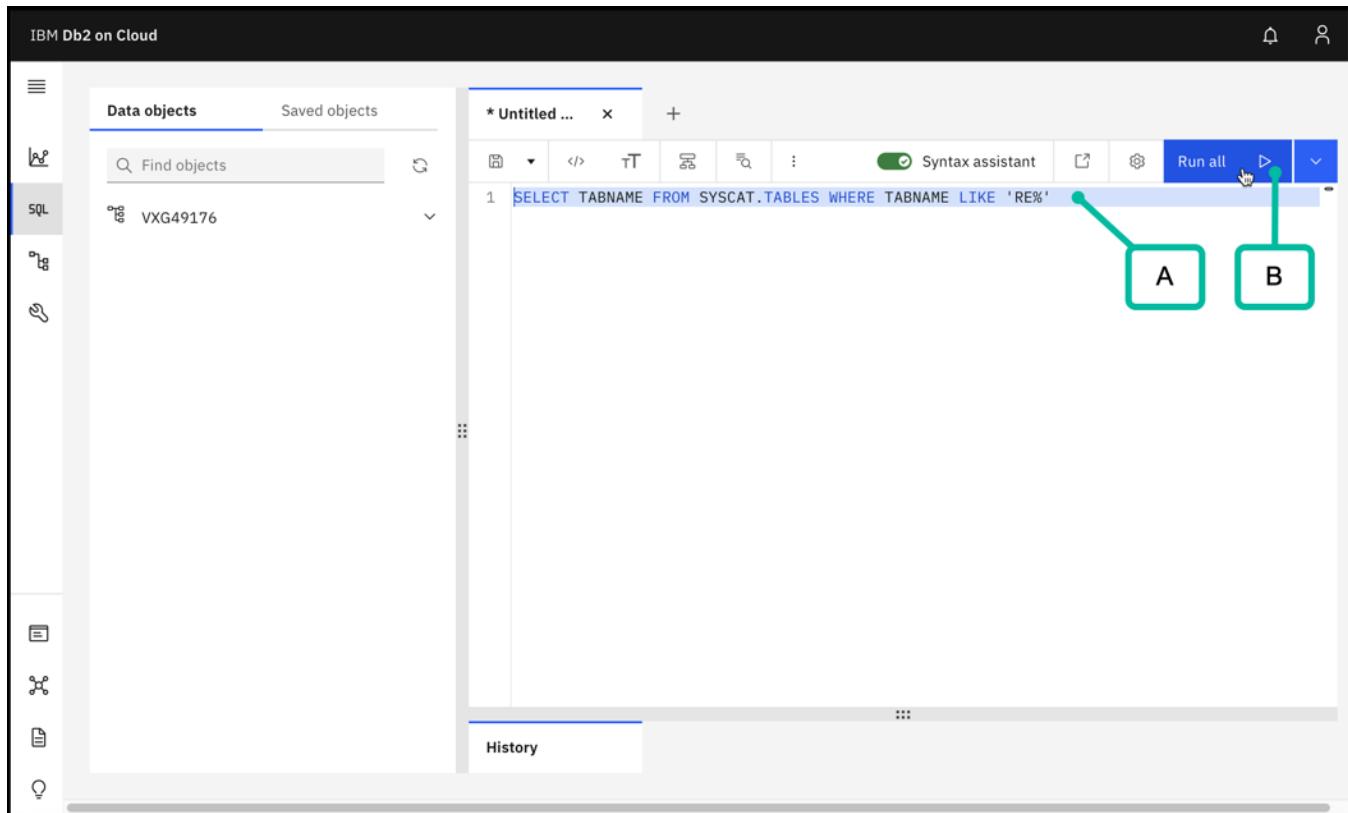
2. When the **Go to UI** button is selected, a new web browser tab named **IBM Db2 on Cloud** is opened. This tab normally consists of an **Overview** screen, which is the default starting point for the Db2 on Cloud User Interface.



\_\_\_\_ 3. Switch to the *Run SQL* screen by clicking the **SQL** icon found in the menu shown on the left-hand side of the screen.



4. [A] Enter the SQL statement **SELECT TABNAME FROM SYSCAT.TABLES WHERE TABNAME LIKE 'RE%'** in the \*Untitled - 1 portion of the *Run SQL* screen. [B] Then, click the **Run all** button to execute the statement. (When executed, this statement will query the database's system catalog and return a list of tables whose name begins with the letters "RE").



```
1 SELECT TABNAME FROM SYSCAT.TABLES WHERE TABNAME LIKE 'RE%'
```

5. When the SQL statement entered in the previous step is executed, a record about its execution will be recorded in a history log, and a summary of the record will be displayed in the **History** window/tab located at the bottom of the screen. If the SQL statement executed successfully, the **Status** column will contain a green circle with a check mark in it; if the statement couldn't be executed, a red circle with a slash in it will appear instead. Find the record for the statement just executed and click on it. (You can find this record by looking for the text of the SQL statement in the **Script** column of the table displayed in the **History** window.)

The screenshot shows the IBM Db2 on Cloud interface. On the left, there's a sidebar with icons for Data objects, Saved objects, Find objects, SQL, History, and other database management tools. The main area has a tab titled '\* Untitled ...' containing the SQL query: 'SELECT TABNAME FROM SYSCAT.TABLES WHERE TABNAME LIKE 'RE%''. Below this is a 'History' section with a table:

Script	Date	Status	Runtime
Untitled - 1	May 18, 2023 5:27:35 PM	✓ 1	0.002 s
SELECT TABNAME FROM SYSCAT.TABLES WHERE TABNAME LIKE ...		✓	0.002 s

6. When you click on the history record for the SQL statement just executed, you will be presented with a new tab named **Results** that contains a table showing the contents of result set produced when the statement was executed, together with the number of records (rows) returned in the result set. (In this example, the text "**Total:4**" indicates that four records were returned.)

The screenshot shows the IBM Db2 on Cloud SQL editor. On the left, there's a sidebar with icons for Data objects, Saved objects, Find objects, SQL, History, Details, Filter table, Total:4, and a trash can icon. The main area has tabs for History and Results, with Results being active. It shows a query window with the SQL command: `SELECT TABNAME FROM SYSCAT.TABLES WHERE TABNAME LIKE 'RE%'`. Below the query window, the results are displayed in a table with one column labeled TABNAME, containing four rows: REFERENCES, REFERENTIAL\_CONSTRAINTS, REF\_CONSTRAINTS, and REG\_VARIABLES.

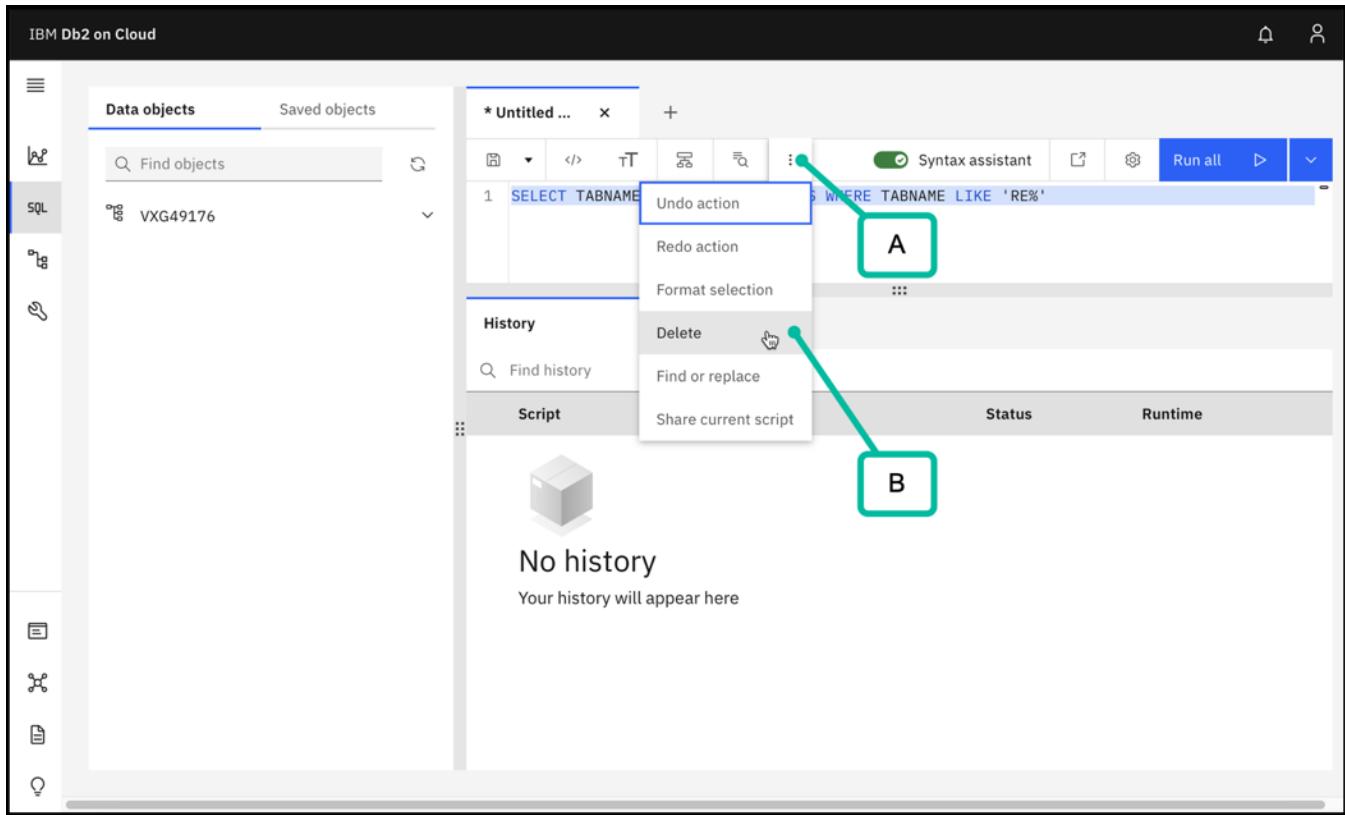
7. Clear the results and the history log by [A] clicking on the **History** tab (to return to the history log display), and then [B] clicking on the trash can icon located at the top left corner of the history log table.

The screenshot shows the IBM Db2 on Cloud interface. On the left is a sidebar with icons for Data objects, Saved objects, Find objects, SQL, and other database management functions. The main area is titled 'Untitled ...' and contains a SQL query: 'SELECT TABNAME FROM SYSCAT.TABLES WHERE TABNAME LIKE 'RE%''. Below the query is a 'History' section. A green callout [A] points to the 'History' tab. Another green callout [B] points to a 'Remove all histories' button with a trash icon.

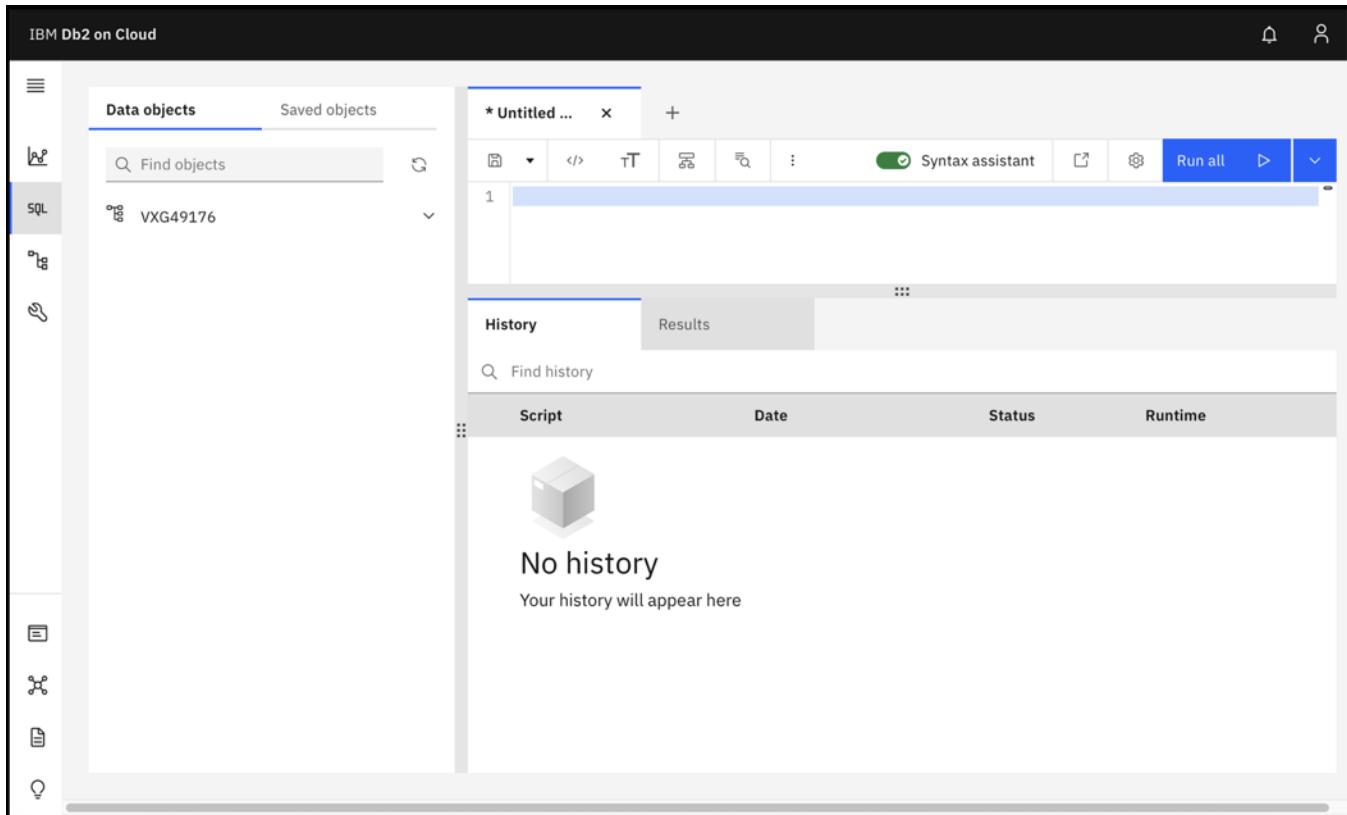
\_\_\_\_ 8. When asked to confirm your decision to delete all results, click the **Yes** button.



\_\_\_\_ 9. Finally, clear the **Untitled-1** section of the *Run SQL* screen by [A] clicking on the “More” icon (the icon consisting of three vertical dots), [B] and then clicking on the **Delete** item shown in the drop-down menu presented.



10. When finished, the *Run SQL* screen should look similar to the way it did when the Db2 on Cloud User Interface was started.



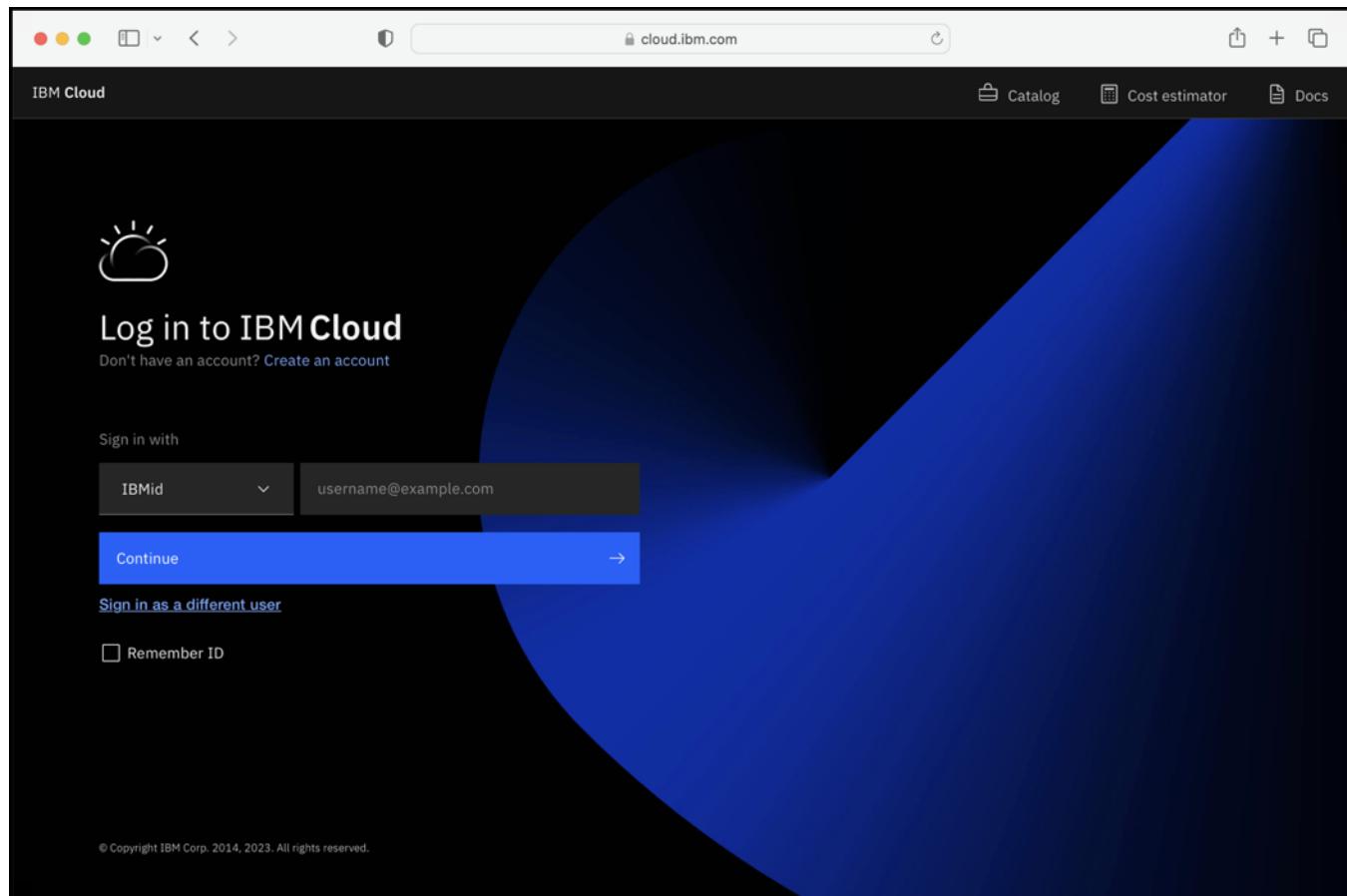
**Congratulations!** You have just created a Db2 on Cloud database and you have seen how to interact with that database using the Db2 on Cloud User Interface. We will use this database with other exercises in this workshop.

### III. Building and populating a Db2 on Cloud database table

When an IBM Db2 on Cloud database is first created, it only contains the system catalog (which we looked at earlier.) For the database to be useful, it must contain at least one user-defined table. So, in this section, we will create a user-defined table and then populate it with data stored in a Comma Separated Values (CSV)-formatted file.

#### 1. Create a table named AIRLINE\_DELAY\_CAUSE.

\_\_\_\_ 1. Open a web browser and go to the *Log in to the IBM Cloud* web site (<https://cloud.ibm.com/login>). You should be presented with a screen that looks something like this:



\_\_\_\_ 2. Log in to IBM Cloud using your IBMid and password. If prompted, log in to IBM as well. Once you are logged in and the *IBM Cloud dashboard* appears, click the hamburger icon (also referred to as the triple bar or trigram symbol) located in the top left corner of the screen. This will open the *Navigation Menu*.

The screenshot shows the IBM Cloud dashboard. The navigation menu on the left is open, with 'Resource list' selected. The main content area displays various service cards: 'Build' (Explore IBM Cloud with this selection of easy starter tutorials and services), 'Monitor your resources' (Get visibility into the performance and health of your resources), 'Visit the IBM Cloud catalog' (Explore our unique product catalog that contains 190+ services and software for your business solutions), 'Virtual Servers' (Deploy your workloads on Virtual Servers with up to 64 vCPU and 512 GB RAM world wide), 'User access' (Enter email addresses below to jump directly into the invite user setup), 'News' (25 IBM Products Win Top Rated Distinction from TrustRadius, IBM Turbonomic SaaS and On-Prem Services Now Available on AWS Marketplace, IBM Spectrum LSF on IBM Cloud: Functional and Performance Updates, IBM Cloud Continuous Delivery Now Supports), and 'Planned maintenance' (View all).

\_\_\_\_ 3. Locate the **Resource list** item in the *Navigation Menu* and click on it. This will take you to the *Resource list* screen, where you can view all of the resources that have been created in your IBM Cloud account.

The screenshot shows the IBM Cloud Resource list screen. The navigation menu on the left is open, with 'Resource list' selected. The main content area displays the same service cards as the dashboard, including 'Build', 'Monitor your resources', 'Visit the IBM Cloud catalog', 'Virtual Servers', 'User access', 'News', and 'Planned maintenance'.

\_\_\_\_ 4. Locate the **Databases** item in the resource list shown and click on the V shown to the left of it. You should see the ID for the Db2 on Cloud database that was created earlier.

The screenshot shows the IBM Cloud Resource list interface. On the left, there's a sidebar with various icons for different service categories like Compute, Containers, Networking, Storage, etc. The main area displays a table with columns: Name, Group, Location, Product, Status, and Tags. There are search and filter bars at the top of the table. The resource list itself is currently collapsed under the 'Databases' category, which has 1+ item. Other collapsed categories include AI / Machine Learning, Analytics, Blockchain, Logging and monitoring, and Migration.

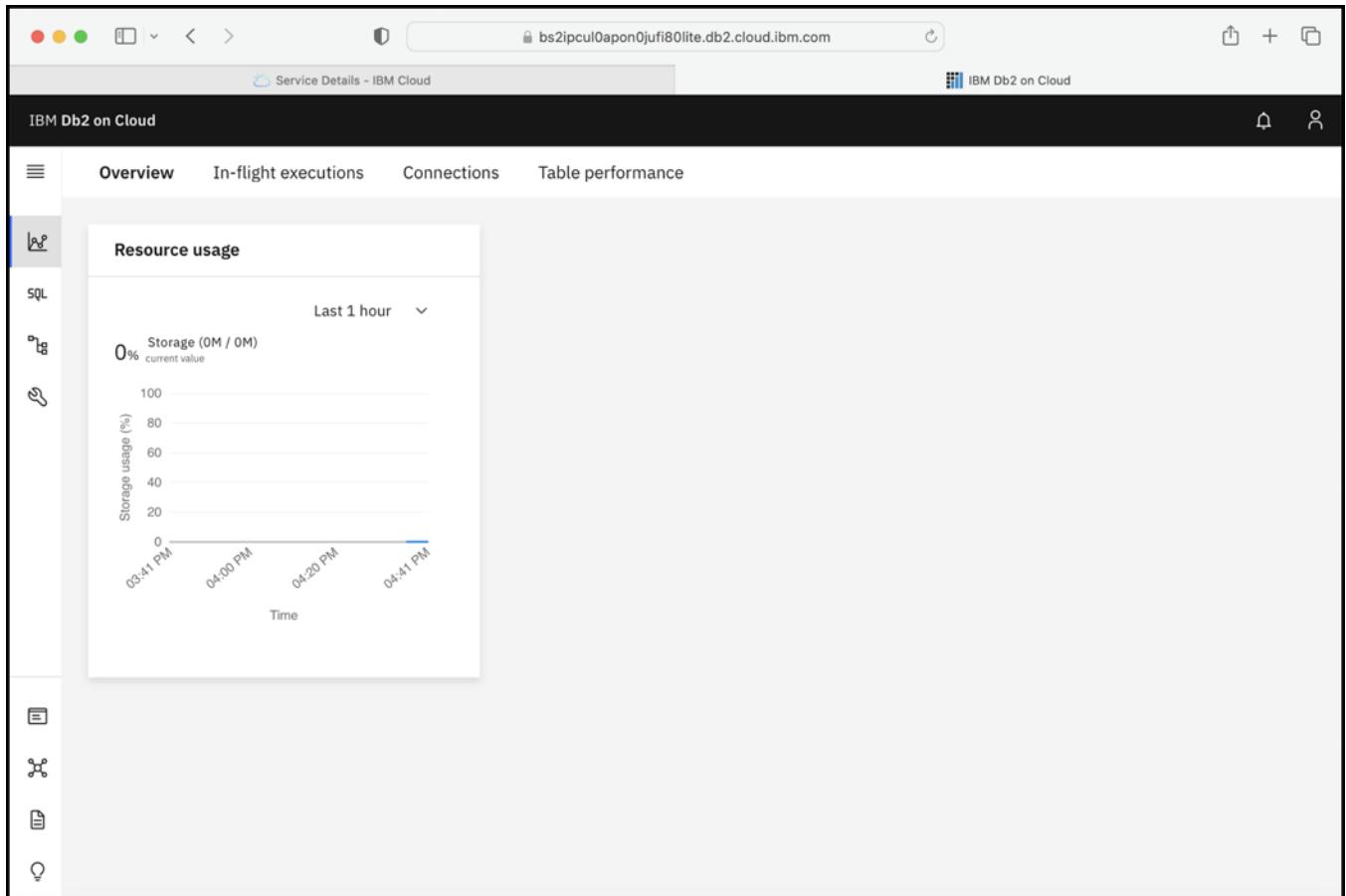
\_\_\_\_ 5. Access the Db2 on Cloud database created earlier by clicking on the database ID shown in the **Resource list**. (In this example, the resource is the Db2 on Cloud database that has the ID *Db2-pc*).

This screenshot provides a detailed view of the Db2-pc database resource. It shows the same sidebar and collapsed categories as the previous screen. In the main table, the 'Databases' row is expanded, revealing the 'Db2-pc' resource. This resource is highlighted with a dark gray background. To its right, there are columns for Default (London), Product (Db2), Status (Active), and a more options button. Below the table, there's a small footer with a message icon.

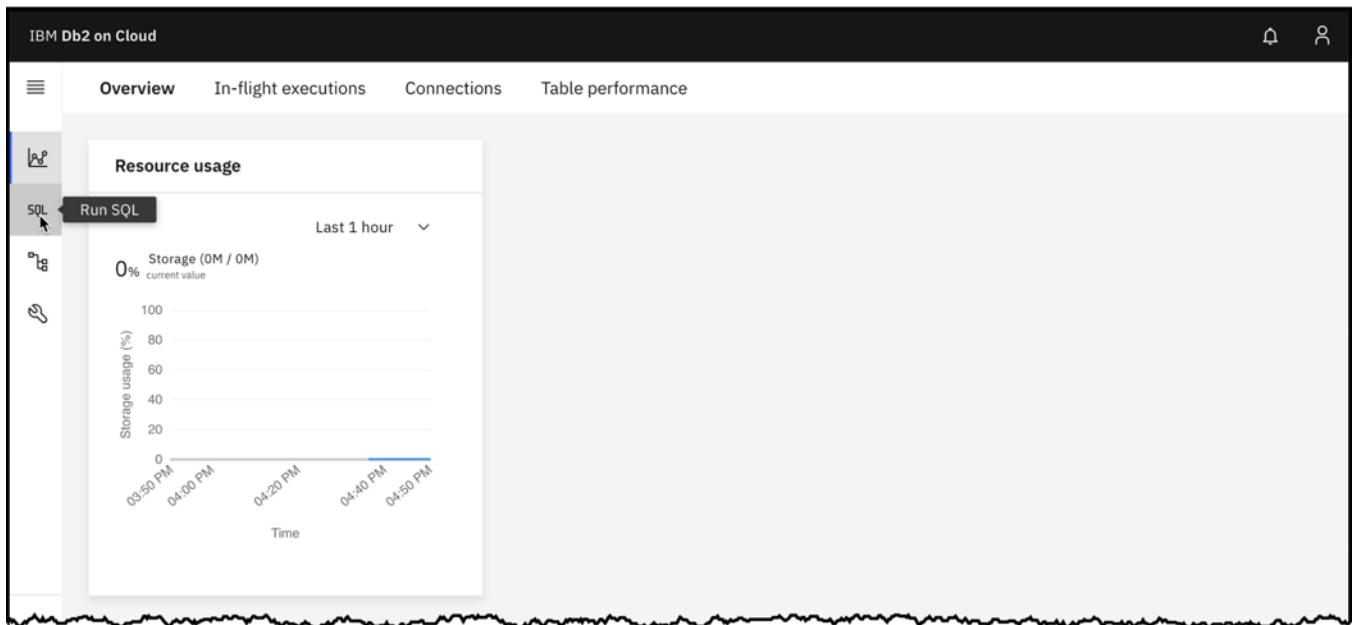
\_\_\_\_ 6. Clicking on a resource ID in the **Resource list** will cause the *Resource list* screen to be replaced with a high-level “management” screen for the resource selected. When the “management” screen for the Db2 on Cloud database resource selected appears, click the **Go to UI** button.

The screenshot shows the IBM Cloud interface for managing a Db2-pc resource. The top navigation bar includes 'IBM Cloud', a search bar, 'Catalog', 'Manage', and user information. Below the navigation is a 'Resource list' section with a single item named 'Db2-pc'. The item status is 'Active' with a green checkmark, and there's an 'Add tags' button. To the right are 'Details' and 'Actions...' buttons. The main content area has three columns: 'Manage' (with links to 'Getting started', 'Service credentials', and 'Connections'), 'Getting started' (with instructions on finding credentials and navigating the UI), and 'Need help?' (with a link to submit a support case). At the bottom are 'Go to UI' (highlighted with a cursor), 'Getting started docs', and 'Support case' buttons.

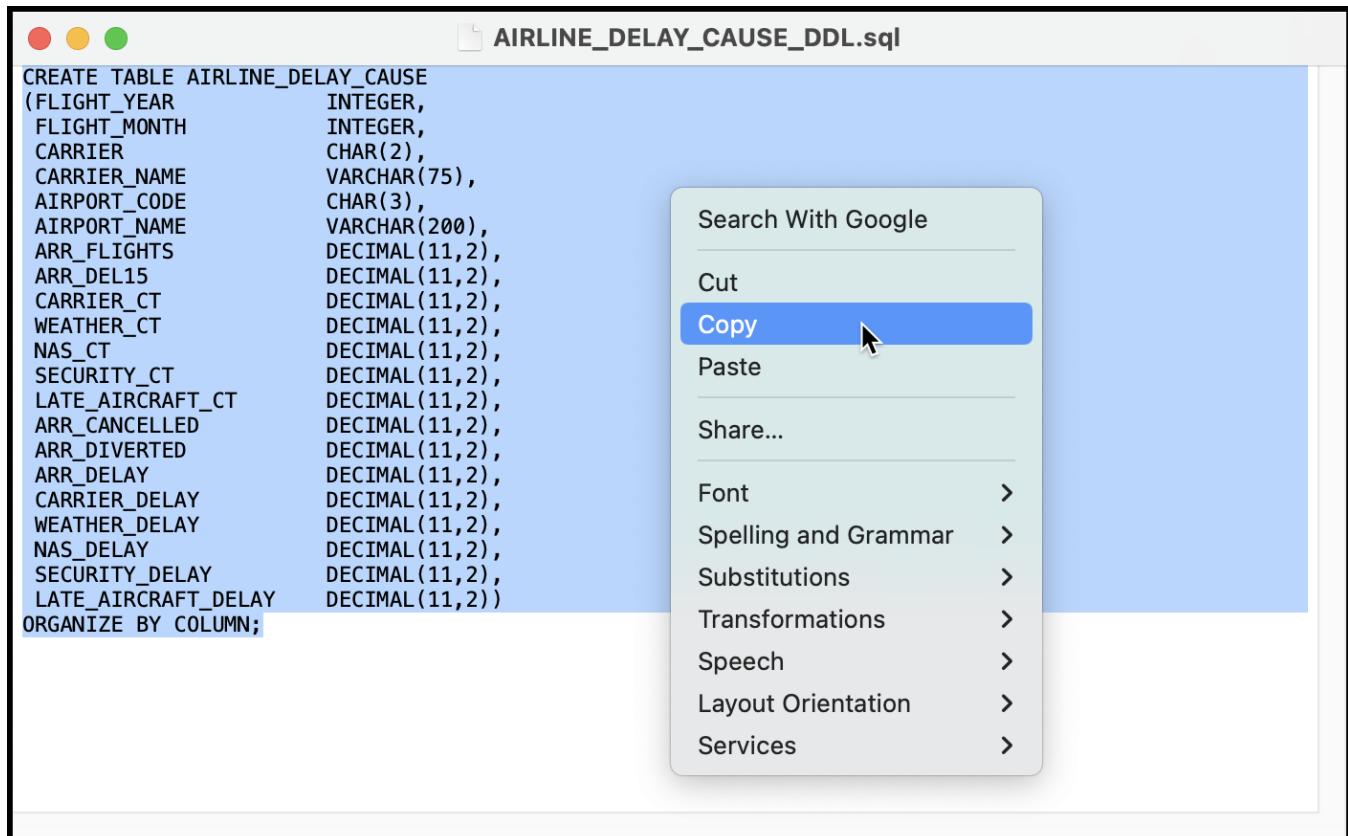
\_\_\_\_ 7. Clicking on the **Go to UI** button will cause a new web browser tab named **IBM Db2 on Cloud** to be opened. This tab usually contains an **Overview** screen, which is the default starting point for the Db2 on Cloud User Interface.



\_\_\_\_ 8. Switch to the *Run SQL* screen by clicking the **SQL** icon found in the menu shown on the left-hand side of the screen.



\_\_\_\_ 9. Minimize your web browser and open the file named **AIRLINE\_DELAY\_CAUSE\_DDL.sql** with a text editor. Then, select everything found in the file and copy it to the clipboard. (In this example, the file was opened with the **TextEdit** application on a MacBook Pro. After everything in the file was selected, a menu was opened by clicking the right button on the mouse. Then, the menu item **Copy** was selected.)



\_\_\_\_ 10. Restore your web browser and [A] use the “hot keys” **command-V** on Mac or **Ctrl-V** on Windows to paste the data copied to the clipboard into the **\*Untitled - 1** window of the *Run SQL* screen. Then, [B] click the **Run all** button to execute the statement just pasted into the window.

The screenshot shows the IBM Db2 on Cloud interface. On the left, there's a sidebar with icons for Data objects, Saved objects, Find objects, and a dropdown menu showing 'VXG49176'. The main area is a code editor titled '\* Untitled ...' containing the following SQL code:

```

1 CREATE TABLE AIRLINE_DELAY_CAUSE
2 (FLIGHT_YEAR           INTEGER,
3 FLIGHT_MONTH          INTEGER,
4 CARRIER                CHAR(2),
5 CARRIER_NAME          VARCHAR(75),
6 AIRPORT_CODE          CHAR(3),
7 AIRPORT_NAME          VARCHAR(200),
8 ARR_FLIGHTS            DECIMAL(11,2),
9 ARR_DEL15              DECIMAL(11,2),
10 CARRIER_CT             DECIMAL(11,2),
11 WEATHER_CT             DECIMAL(11,2),
12 NAS_CT                 DECIMAL(11,2),
13 SECURITY_CT             DECIMAL(11,2),
14 LATE_AIRCRAFT_CT       DECIMAL(11,2),
15 ARR_CANCELLED          DECIMAL(11,2),
16 ARR_DIVERTED           DECIMAL(11,2),
17 CARRIER_DELAY           DECIMAL(11,2),
18 WEATHER_DELAY           DECIMAL(11,2),
19 NAS_DELAY               DECIMAL(11,2),
20 SECURITY_DELAY          DECIMAL(11,2),
21 LATE_AIRCRAFT_DELAY     DECIMAL(11,2))
22 ORGANIZE BY COLUMN;
23

```

At the top right of the code editor, there are several buttons: Syntax assistant, Run all (highlighted with a green arrow), and others. A green callout box labeled 'A' points to the 'Run all' button. Another green callout box labeled 'B' points to the 'Syntax assistant' button.

11. Check the **History** window/tab located at the bottom of the screen. If the **Status** column for the record containing the text “**CREATE TABLE AIRLINE\_DELAY\_CAUSE ...**” has a green circle with a check mark in it, a table named **AIRLINE\_DELAY\_CAUSE** was successfully created. If an error occurred, clear the contents of the **\*Untitled - 1** window and redo steps 9 and 10 (make sure you copy every line in the file **AIRLINE\_DELAY\_CAUSE\_DDL.sql**.)

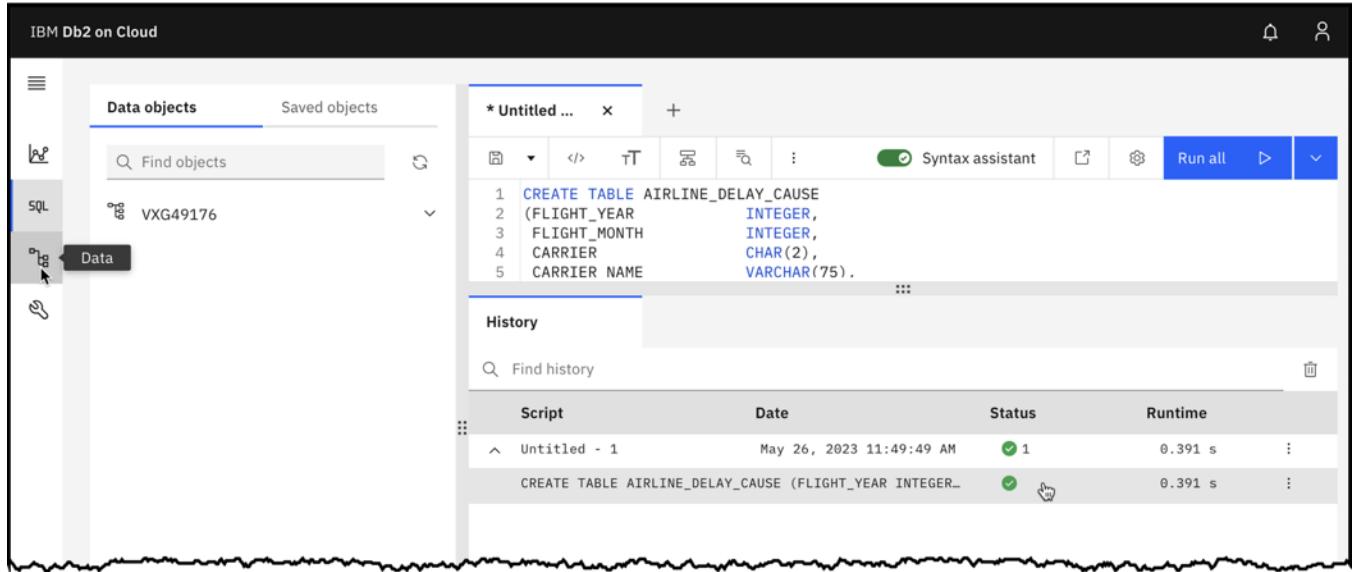
The screenshot shows the IBM Db2 on Cloud interface with the History window open at the bottom. The history table contains the following entries:

Script	Date	Status	Runtime
Untitled - 1	May 26, 2023 11:49:49 AM	<span style="color: green;">✓ 1</span>	0.391 s
CREATE TABLE AIRLINE_DELAY_CAUSE (FLIGHT_YEAR INTEGER...		<span style="color: green;">✓</span>	0.391 s

The status column for the second entry shows a green circle with a checkmark, indicating success. The runtime for both entries is 0.391 seconds.

## 2. Populate the AIRLINE\_DELAY\_CAUSE table.

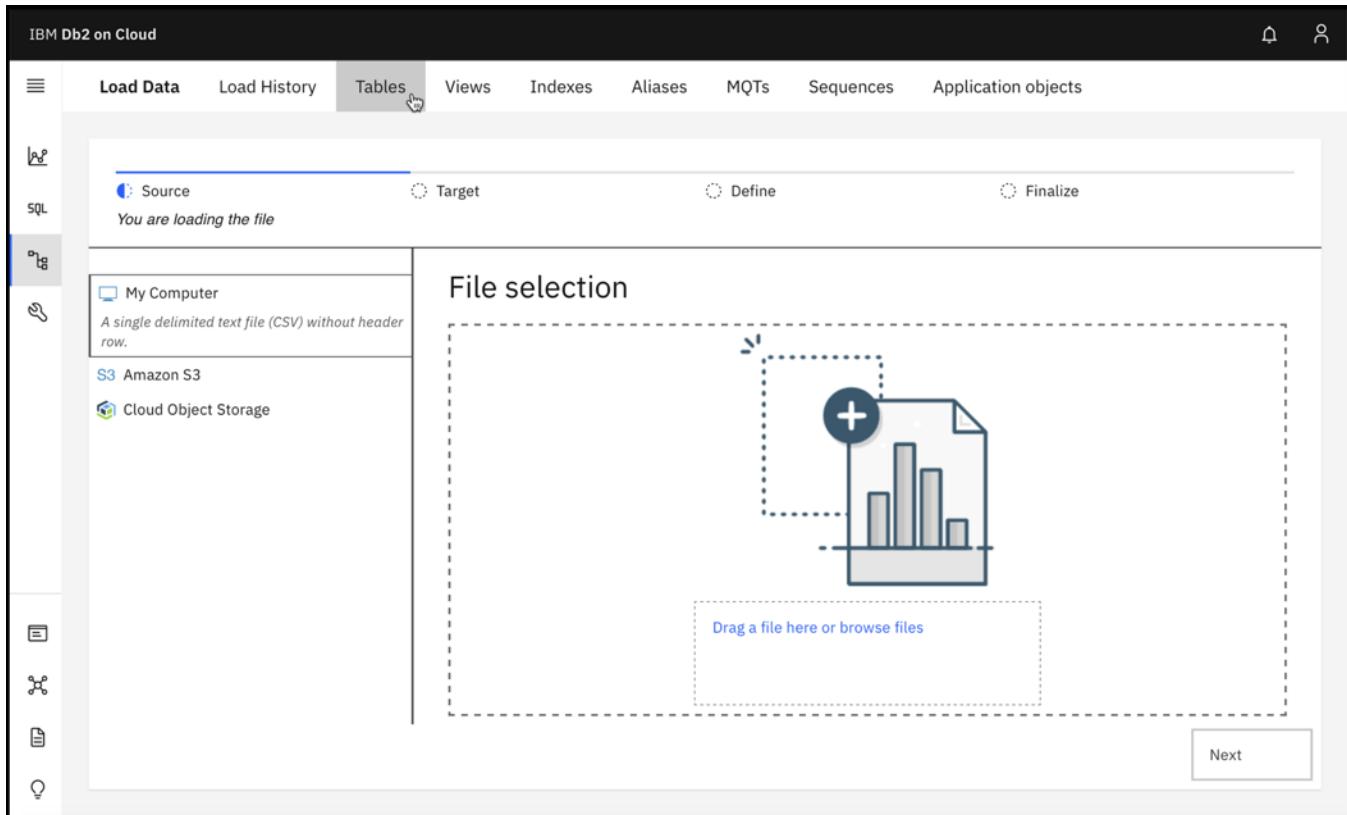
\_\_\_\_ 1. Switch to the *Data* screen by clicking the **Data** icon (located immediately below the **SQL** icon) found in the menu shown on the left-hand side of the screen.



```
CREATE TABLE AIRLINE_DELAY_CAUSE
(FLIGHT_YEAR      INTEGER,
FLIGHT_MONTH     INTEGER,
CARRIER          CHAR(2),
CARRIER_NAME     VARCHAR(75).
```

Script	Date	Status	Runtime
CREATE TABLE AIRLINE_DELAY_CAUSE (FLIGHT_YEAR INTEGER...	May 26, 2023 11:49:49 AM	✓ 1	0.391 s

\_\_\_\_ 2. When you go to the *Data* screen, the first thing you see is the *Load Data* window. Before we use this window, let's confirm that the **AIRLINE\_DELAY\_CAUSE** table we just created exists. Click on the **Tables** tab to go to the *Tables* screen.



3. On the *Tables* screen, locate the **Schema Name** that is associated with your IBMid and click on the check-box beside it. (In this example, the schema name is **VXG49176**.) Because we are using the IBM Db2 on Cloud “Lite” plan, there should be only one schema name in the list.



4. Once you have selected a schema in the *Schemas* window, a list of tables that have been defined for that schema will be displayed in a *Tables* window that is opened on the right-hand side of the screen. You should see the **AIRLINE\_DELAY\_CAUSE** table in the list shown in this window.

IBM Db2 on Cloud

Load Data Load History Tables Views Indexes Aliases MQTs Sequences Application objects

Find schemas or tables Refresh

Schemas

Name	Definer type	Tables ▲
VXG49176	User	1

Tables

Name ▾	Schema	Properties
AIRLINE_DELAY...	VXG49176	...

5. Click on the link for the **AIRLINE\_DELAY\_CAUSE** table in the *Tables* window. You should see detailed information about the table, such as column names and data types that were provided when the table was created.

IBM Db2 on Cloud

Load Data Load History Tables Views Indexes Aliases MQTs Sequences Application objects

Find schemas or tables Refresh

Tables

Name ▾	Schema	Properties
AIRLINE_DELAY...	VXG49176	...

Table definition

AIRLINE\_DELAY\_CAUSE

No statistics available.

Name	Data type	Nullable	Length	Scale
FLIGHT_YEAR	INTEGER	Y		0
FLIGHT_MONTH	INTEGER	Y		0
CARRIER	CHAR	Y	2	0
CARRIER_NAME	VARCHAR	Y	75	0
AIRPORT_CODE	CHAR	Y	3	0
AIRPORT_NAME	VARCHAR	Y	200	0
ARR_FLIGHTS	DECIMAL	Y	11	2

Total: 1, selected: 0

View data

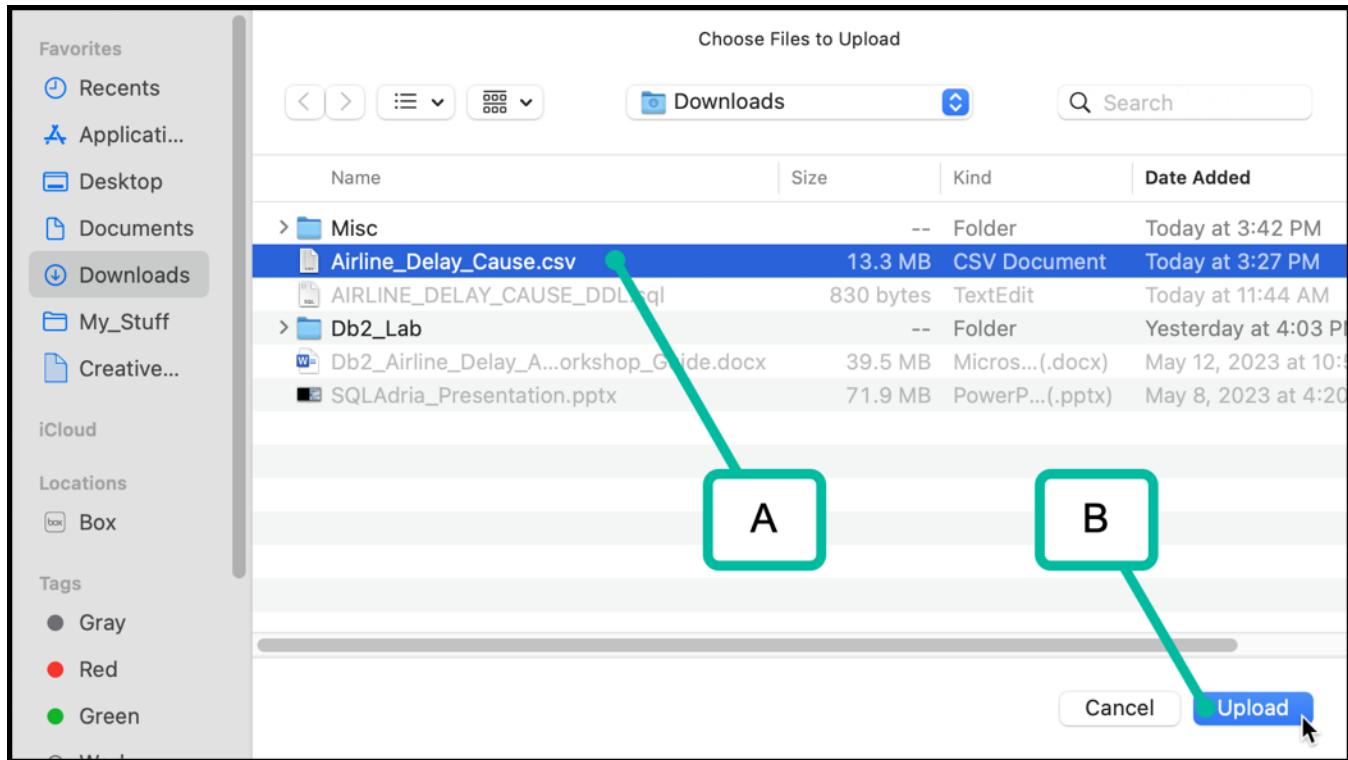
6. Go back to the *Load Data* screen by clicking on the **Load Data** tab.

The screenshot shows the IBM Db2 on Cloud service details page. The top navigation bar includes 'Load Data', 'Load History', 'Tables', 'Views', 'Indexes', 'Aliases', 'MQTs', 'Sequences', and 'Application objects'. The 'Tables' tab is selected. On the left, there's a sidebar with 'Tables' and 'Schemas' tabs. The main area shows a table definition for 'ALINE\_DELAY\_CAUSE' with two columns: 'FLIGHT\_YEAR' (INTEGER) and 'FLIGHT\_MONTH' (INTEGER). Below the table definition, there's a table with columns 'Name', 'Schema', and 'Properties', showing one entry: 'ALINE\_DELAY...' under 'Name' and 'VXG49176' under 'Schema'.

7. Once you return to the *Load Data* screen, click on the **Drag a file here or browse files** link to open a *Finder* window on Mac or a *File Explorer* window on Windows.

The screenshot shows the 'Load Data' screen in IBM Db2 on Cloud. The 'Source' tab is selected. The main area is titled 'File selection' and contains a large dashed box with a central document icon featuring a bar chart. Below the document icon is a blue rectangular box containing the text 'Drag a file here or browse files'. There are also four radio buttons labeled 'Source', 'Target', 'Define', and 'Finalize'. To the left, there's a sidebar with various icons and a list of storage options: 'My Computer' (selected), 'Amazon S3', and 'Cloud Object Storage'. At the bottom right of the main area is a 'Next' button.

8. [A] Select the file named **Airline\_Delay\_Cause.csv** and [B] click the **Upload** button.



9. Verify that the file name **Airline\_Delay\_Cause.csv** appears under the **Selected file** heading on the *Load Data* screen. Then, click the **Next** button located in the bottom right-hand corner of the screen.

\_\_\_\_ 10. From the *Select a load target* screen, [A] click on the *Schema* name associated with your IBMid. (In this example, the schema name is **VXG49176**.) Then, [B] click on the **AIRLINE\_DELAY\_CAUSE** table shown in the *Table* list. Next, [C] select the **Overwrite table with new data** radio button. Finally, [D] click the **Next** button shown in the bottom right-hand corner of the screen.

IBM Db2 on Cloud

Load Data Load History Tables Views Indexes Aliases MQTs Sequences Application objects

Source Target Define Finalize

You are loading the file *Airline\_Delay\_Cause.csv* into **VXG49176.AIRLINE\_DELAY\_CAUSE**

Select a load target

Schema

Find schemas

VXG49176 A

Table

New table +

Find tables in VXG49176

AIRLINE\_DELAY\_CAUSE B

Table definition

**AIRLINE\_DELAY\_CAUSE**  
Approximate -1.0 rows (-1.0 KB)  
No statistics available

overwrite\_option

Append new data  
 Overwrite table with new data C

All existing data will be deleted from the table whether or not the loading action completes successfully.

COLUMN DATA TYPE NULLABLE NAME

Back Next D

\_\_\_\_ 11. Verify that the **Airline\_Delay\_Cause.csv** file used contains valid data. Then, [A] make sure the **Header in first row** switch is on (green) and [B] click the **Next** button.

The screenshot shows the IBM Db2 on Cloud interface for loading data. The main area displays a preview of a CSV file named 'Airline\_Delay\_Cause.csv' into a target table 'VXG49176.AIRLINE\_DELAY\_CAUSE'. The preview shows 10 rows of data with columns: FLIGHT\_YEAR, FLIGHT\_MONTH, CARRIER, and CARRIER\_NAME. The CARRIER column contains the value '9E' for all rows. The CARRIER\_NAME column contains the value 'Endeavor Air Inc.' for all rows.

Annotations:

- Annotation A points to the 'Header in first row:' checkbox, which is checked.
- Annotation B points to the 'Time & date format:' dropdown, which is currently set to 'None'.
- A green arrow points from annotation A to the 'Header in first row:' checkbox.
- A green arrow points from annotation B to the 'Time & date format:' dropdown.
- The 'Next' button at the bottom right is highlighted with a blue box and a mouse cursor icon.

	FLIGHT_YEAR INTEGER	FLIGHT_MONTH INTEGER	CARRIER CHARACTER	CARRIER_NAME VARCHAR
1	2022	12	9E	Endeavor Air Inc.
2	2022	12	9E	Endeavor Air Inc.
3	2022	12	9E	Endeavor Air Inc.
4	2022	12	9E	Endeavor Air Inc.
5	2022	12	9E	Endeavor Air Inc.
6	2022	12	9E	Endeavor Air Inc.
7	2022	12	9E	Endeavor Air Inc.
8	2022	12	9E	Endeavor Air Inc.
9	2022	12	9E	Endeavor Air Inc.
10	2022	12	9E	Endeavor Air Inc.

12. Look over the *Review settings* screen to ensure everything is correct. Then, start the data load process by clicking the **Begin Load** button.

The screenshot shows the IBM Db2 on Cloud interface with the 'Load Data' tab selected. The main area displays a 'Review settings' screen for loading data from 'Airline\_Delay\_Cause.csv' into 'VXG49176.AIRLINE\_DELAY\_CAUSE'. The 'Source' tab is selected. The 'Option' section includes a setting for 'Maximum number of warnings' set to '1000'. At the bottom right are 'Back' and 'Begin Load' buttons, with the 'Begin Load' button being highlighted.

13. When the **Begin Load** button is selected, the *Review settings* screen will be replaced by the *Load details* screen, and you will receive a notification that there is a **Load in progress**. This notification will include information about where data is coming from (in this example, the **Airline\_Delay\_Cause.csv** file stored on **My Computer**) and where it is going (the **VXG49176.AIRLINE\_DELAY\_CAUSE** table).

The screenshot shows the IBM Db2 on Cloud interface with the 'Load Data' tab selected. In the top right corner, there is a notifications icon with a red '1' indicating one notification. The main area is titled 'Load details' and shows a 'LOADING' status. It displays the source file 'Airline\_Delay\_Cause.csv' and the target table 'VXG49176.AIRLINE\_DELAY\_CAUSE'. Below this, the 'Status' section shows 'Loading' with counts of 0 rows read, 0 rows loaded, and 0 rows rejected. It also shows the start time as '05/26/2023 3:49:18 PM' and the elapsed time as 'a few seconds ago'. To the right, a 'Did you know?' box provides information about exporting SQL statements. A vertical sidebar on the left contains icons for navigation and management.

14. Once the load operation is completed, the *Load details* screen will be updated to reflect the number of rows read from the source file, the number of rows loaded into the table specified, and the number of rows rejected. You will also receive a notification that **The data load job succeeded** (or failed).

The screenshot shows the IBM Db2 on Cloud interface for managing data loads. The top navigation bar includes options like Load Data, Load History, Tables, Views, Indexes, Aliases, MQTs, Sequences, Notifications, and Clear all. A notification bar at the top right indicates '1' notifications, with a message: 'The data load job succeeded. Load Airline\_Delay\_Cause.csv from My Computer to VXG49176.AIRLINE\_DELAY\_CAUSE' at 2023/05/26, 03:49 PM, with a 'View details' link.

The main area is titled 'Load details' and shows a summary of the completed load:

- My computer: Airline\_Delay\_Cause.csv
- Target: VXG49176.AIRLINE\_DELAY\_CAUSE
- Status: COMPLETE (indicated by a green checkmark icon)
- Rows read: 80,104
- Rows loaded: 80,104
- Rows rejected: 0

Timing information:

- Start time: 05/26/2023 3:49:18 PM
- End time: 05/26/2023 3:49:58 PM

On the right side, there are sections for 'Errors' (0) and 'Warnings' (0), both of which are currently empty. A large blue circular progress indicator is present in the center of the page.

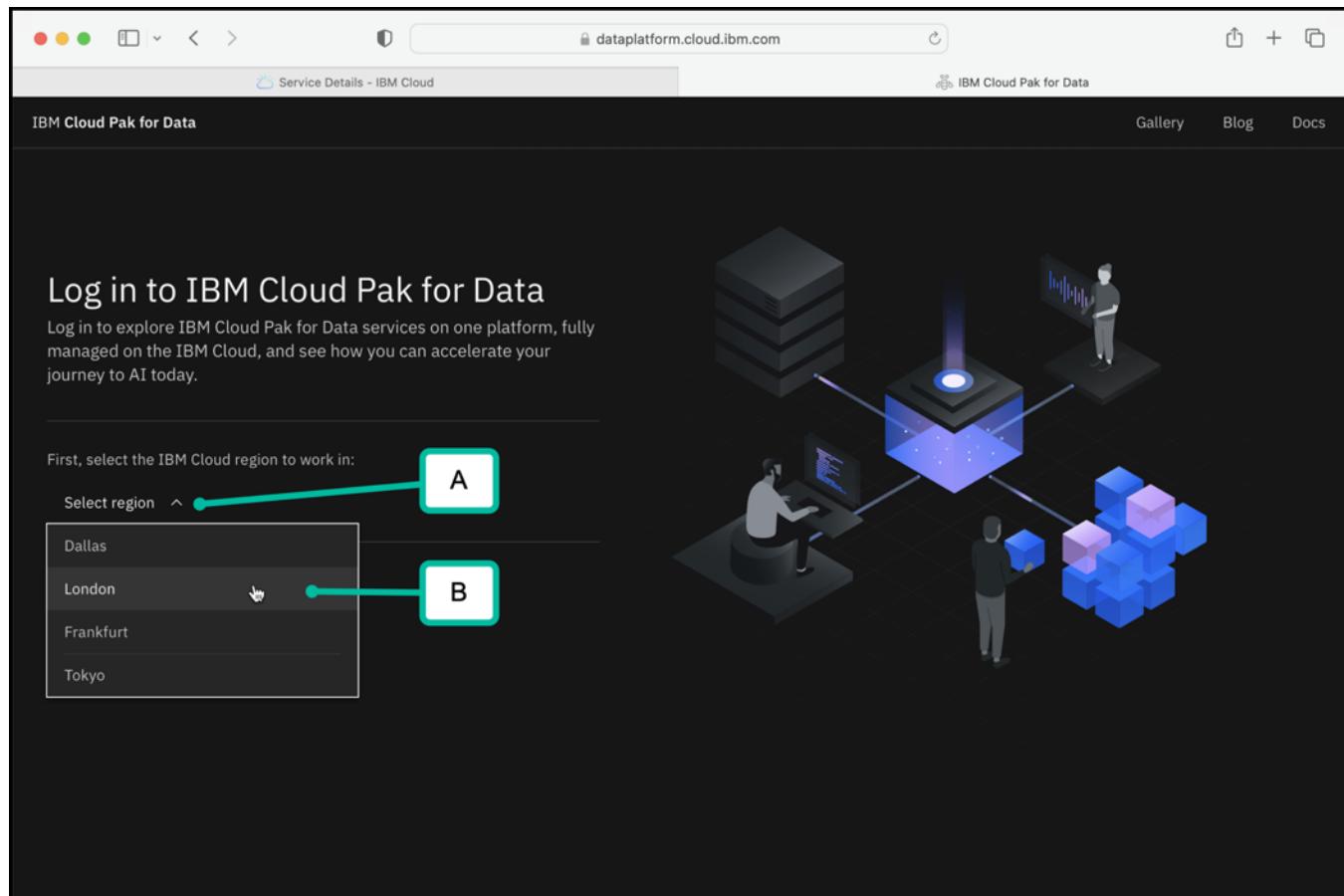
**Congratulations!** You have just created and populated a table named **AIRLINE\_DELAY\_CAUSE** in a Db2 on Cloud database. We will use the data in this table in another exercise a little later.

## IV. Using IBM Watson Studio to run a Jupyter Notebook

The free version of IBM Watson Studio (offered with IBM Cloud Pak for Data as a Service) will be used to execute the Jupyter Notebooks that are needed to complete the remaining exercises in this workshop. So, in this section you will set up a free IBM Cloud Pak for Data as a Service account and then use that account to run a simple Jupyter Notebook.

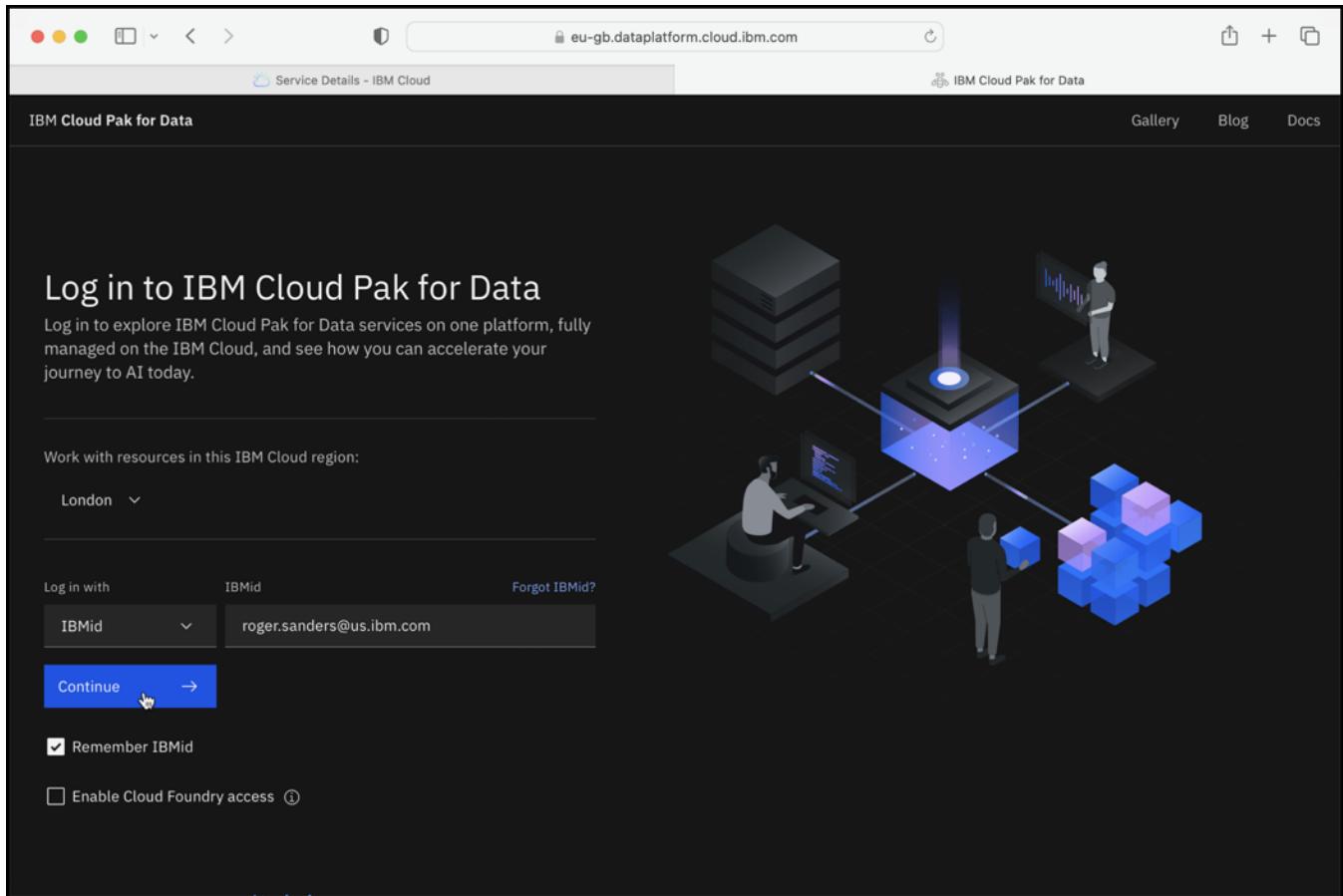
### 1. Create an IBM Cloud Pak for Data as a Service account.

\_\_\_\_ 1. Open a web browser and go to the *Log in to IBM Cloud Pak for Data* web site (<https://dataplateform.cloud.ibm.com/login>). Once there, [A] click on the **Select region** button to display a dropdown list of locations where Cloud Pak for Data as a Service can be deployed. Then, [B] choose the location where your IBM Cloud Pak as a Service account is to be provisioned. (In this example, the **London** location was selected.)



\_\_\_\_ 2. When the log in fields appear, make sure the **Log in with** field contains the value **IDBId**. Then, enter the email address you used to create your IBMid in the **IBMid** field (located beside the **Log in with** field).

**IMPORTANT: Please use the same IBMid that was used earlier to create an IBM Cloud account and provision a Db2 on Cloud Lite plan database.** When finished, click the **Continue** button.



3. When the *Provide your information to continue* screen appears, [A] enter your company name in the **Company name** field and [B] your telephone number in the **Phone number** field. (Click on the **Phone number** dropdown list and select the correct country code if your country code is not +1.) Then, [C] indicate how you would like for IBM to contact you by selecting the **by email** and/or the **by phone** check-boxes. Finally, [D] click the **Continue** button.

The screenshot shows a web browser window with the URL [eu-gb.dataplatform.cloud.ibm.com](https://eu-gb.dataplatform.cloud.ibm.com). The page title is "Service Details - IBM Cloud". The main content is titled "IBM Cloud Pak for Data" and asks for information to continue. It includes fields for "Company name" (with "IBM" entered) and "Phone number" (+1 9999999999). There is a "Continue" button with a cursor icon. To the right, there is a marketing consent section with two checkboxes: "by email" (checked) and "by phone" (unchecked). Below this is a note about marketing consent and a link to the "IBM Privacy Statement". At the bottom, there is a statement about accepting the "IBM Privacy Statement and Terms and Conditions". A "Cookie Preferences" button is visible in the bottom right corner.

Provide your information to continue

Company name  
IBM

Phone number  
+1 9999999999

Continue

IBM may use my contact data to keep me informed of products, services, and offerings:

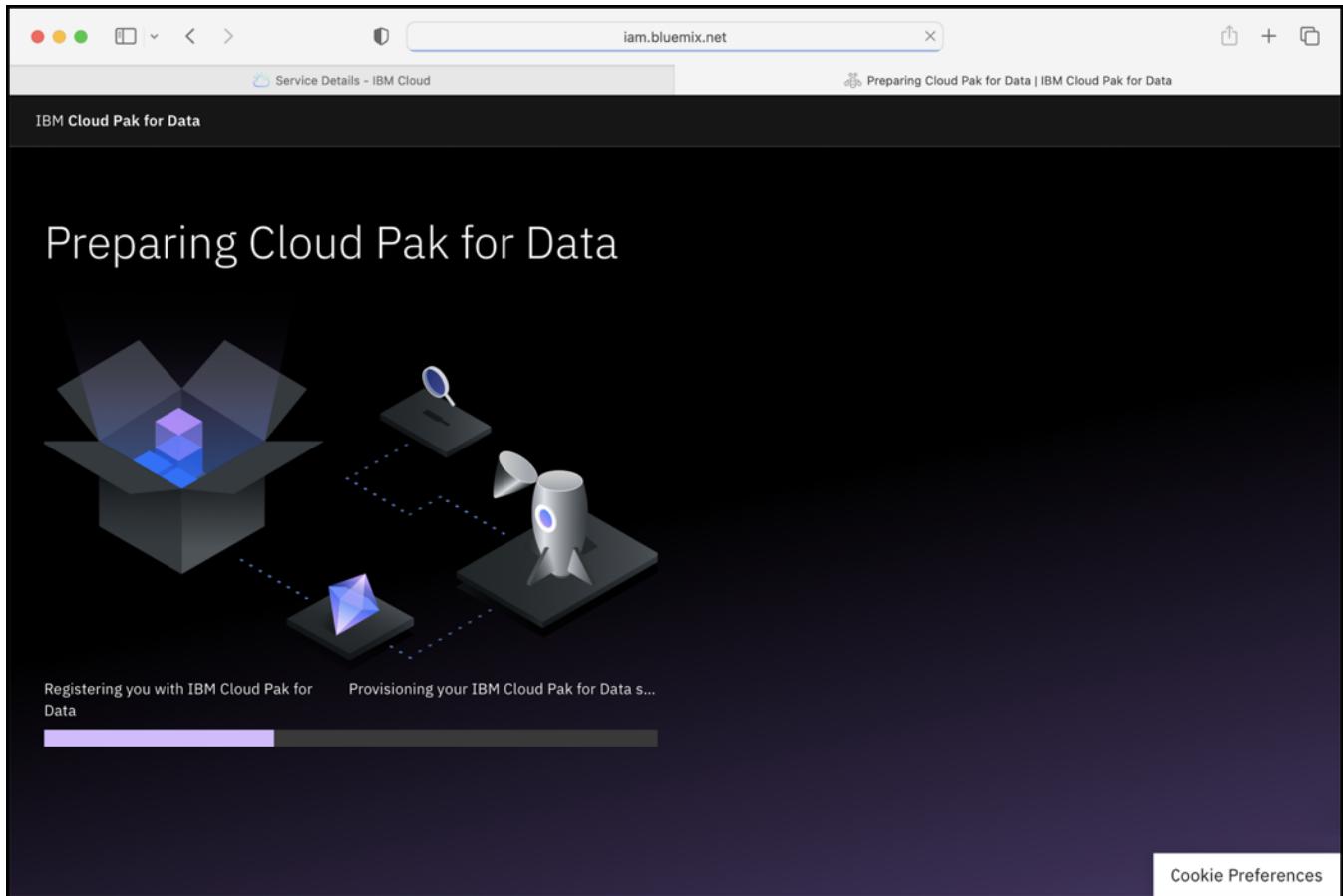
by email       by phone

You can withdraw your marketing consent at any time by submitting an [opt-out request](#). Also, you may unsubscribe from receiving marketing emails by clicking the unsubscribe link in each email. More information on our processing can be found in the [IBM Privacy Statement](#).

By submitting this form, I acknowledge that I have read and understand the IBM Privacy Statement and I accept the product Terms and Conditions of this registration form.

Cookie Preferences

4. While you are being registered with IBM Cloud Pak for Data and your IBM Cloud Pak for Data as a Service environment is being created, you should see a *Preparing Cloud Pak for Data* screen similar to this:



\_\_\_\_ 5. Once your IBM Cloud Pak for Data as a Service environment has been created, you should see a *Welcome* screen that looks something like this:

Welcome, Roger!

**Take a tutorial**  
Step through implementing a Data fabric use case in a sample project.

**Work with data**  
Create a project for your team to prepare data, find insights, or build models.

**Find what you need**  
Open this interactive map to find the tools that you need for your tasks.

**Quick start**

- Build customer profiles** with IBM Match 360 with Watson
- Catalog and govern data** with Watson Knowledge Catalog
- Build and manage ML models** with Watson Studio
- Query data anywhere** with Watson Query

**Projects** +

**No recent projects**  
After you create projects, you'll see your recently updated projects here.  
New project +

**Catalogs** +

**No catalogs**  
Your catalogs show here after you create or join them. Click New catalog to get started.  
New catalog +

**Notifications**

**No notifications**  
You will see your most recent notifications here.

**Deployments** ⓘ +

**New in gallery**

Feedback

**Congratulations!** You now have an active **IBM Cloud Pak for Data as a Service** account that can be used to perform the rest of the exercises in this workshop.

## 2. Create a new project.

1. Create a new project by clicking on the + icon located in the top right-hand corner of the **Projects** box. This will cause the IBM Cloud Pak for Data as a Service *Welcome* screen to be replaced with a *Create a project* screen.

Welcome, Roger!

**Take a tutorial**  
Step through implementing a Data fabric use case in a sample project.

**Work with data**  
Create a project for your team to prepare data, find insights, or build models.

**Find what you need**  
Open this interactive map to find the tools that you need for your tasks.

**Quick start**

- Build customer profiles with IBM Match 360 with Watson
- Catalog and govern data with Watson Knowledge Catalog
- Build and manage ML models with Watson Studio
- Query data anywhere with Watson Query

**Create new project**

**Projects**

No recent projects  
After you create projects, you'll see your recently updated projects here.  
New project +

**Catalogs**

No catalogs  
Your catalogs show here after you create or join them. Click New catalog to get started.  
New catalog +

**Notifications**

No notifications  
You will see your most recent notifications here.

**Deployments**

New in gallery

2. When the *Create a project* screen is displayed, click on the **Create an empty project** tile.

**Create a project**

Choose whether to create an empty project or to preload your project with data and analytical assets. Add collaborators and data, and then choose the right tools to accomplish your goals. Add services as necessary.

**Create an empty project**

Add the data you want to prepare, analyze, or model. Choose tools based on how you want to work: write code, create a flow on a graphical canvas, or automatically build models.

**Create a project from a sample or file**

Get started fast by loading existing assets. Choose a project file from your system, or choose a curated sample project.

**USE TO**

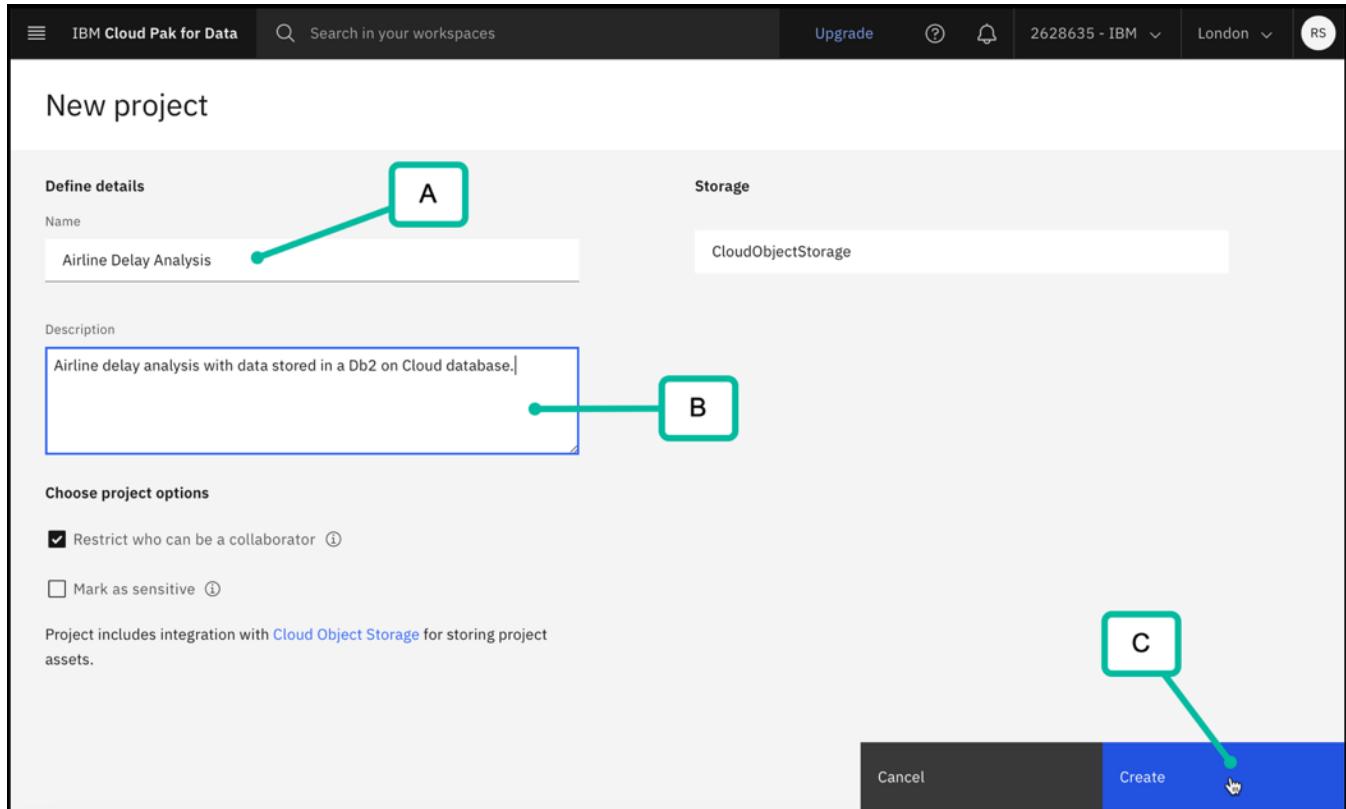
**Create an empty project**

- Prepare and visualize data
- Analyze data in notebooks
- Train models

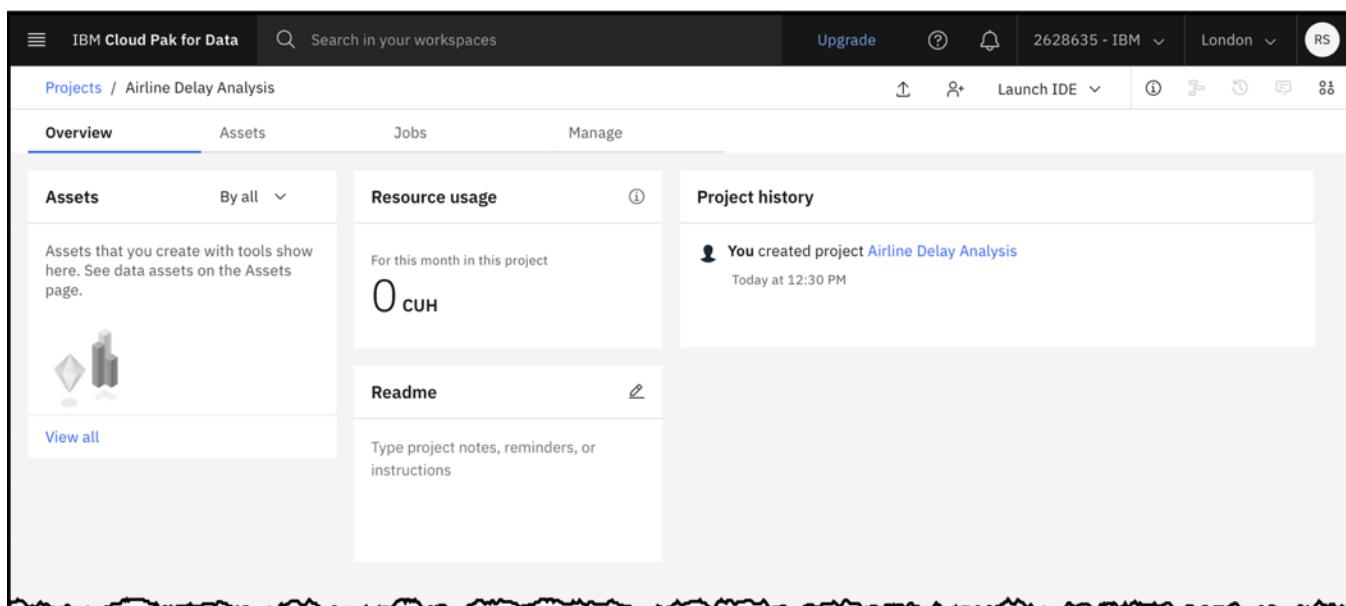
**Create a project from a sample or file**

- Learn by example
- Build on existing work
- Run tutorials

\_\_\_\_ 3. When the *New project* screen is presented, [A] enter a project name in the **Name** field. Then, [B] enter a project description in the **Description** field (*optional*). In this example, the project name “*Airline Delay Analysis*” and the description “*Airline delay analysis with data stored in a Db2 on Cloud database.*” were used. Finally, [C] click the **Create** button.

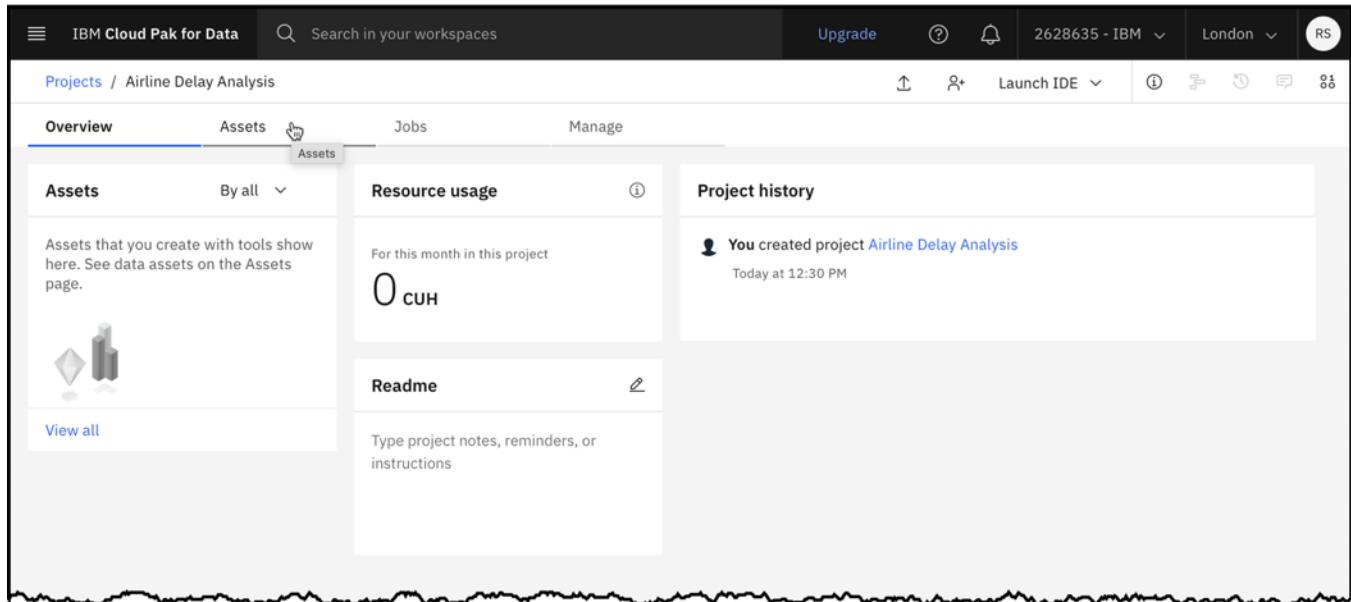


\_\_\_\_ 4. Once the project is created, you should see a *Project Overview* screen that looks something like this:



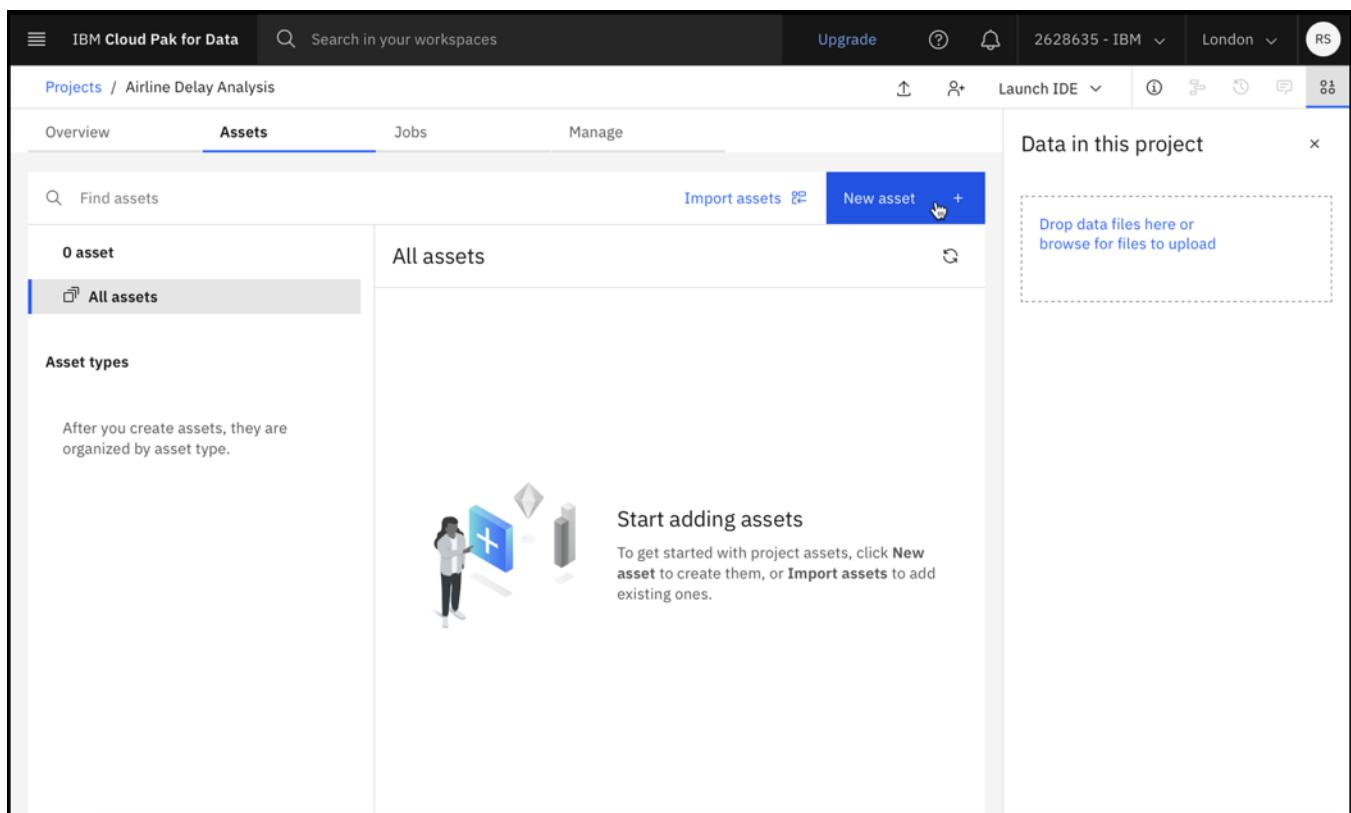
### 3. Add a Jupyter Notebook to the project.

1. From the *Project Overview* screen, click on the **Assets** tab.



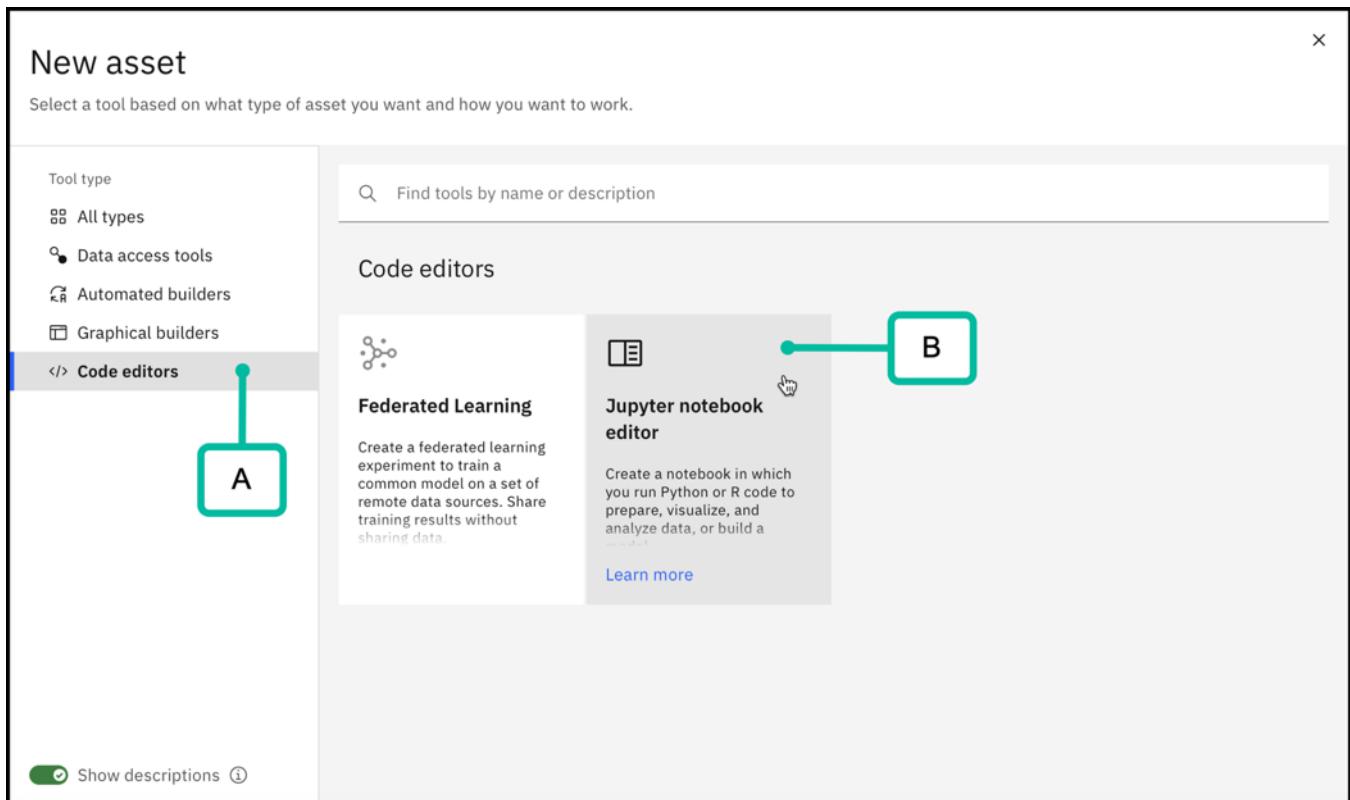
The screenshot shows the 'Project Overview' screen for the 'Airline Delay Analysis' project. The top navigation bar includes 'IBM Cloud Pak for Data', a search bar, 'Upgrade', 'Help', '2628635 - IBM', 'London', and a user icon. The main area has tabs for 'Overview', 'Assets' (which is highlighted with a blue underline), 'Jobs', and 'Manage'. The 'Assets' section displays a message: 'Assets that you create with tools show here. See data assets on the Assets page.' It features a bar chart icon and a 'View all' link. The 'Resource usage' section shows '0 cuH' for the current month. The 'Project history' section shows a message: 'You created project Airline Delay Analysis Today at 12:30 PM'.

2. When the Assets screen is displayed, click on the **New asset +** button.

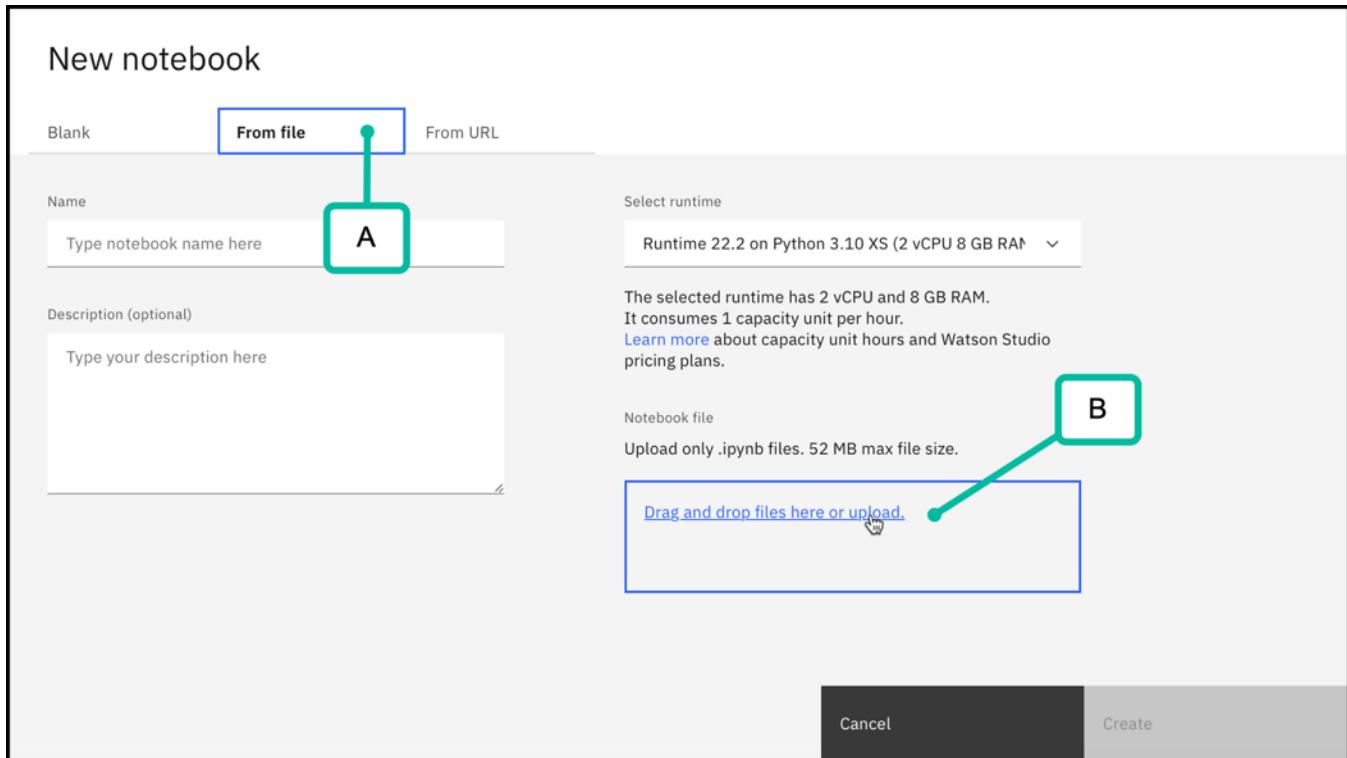


The screenshot shows the 'Assets' screen for the 'Airline Delay Analysis' project. The top navigation bar is identical to the previous screenshot. The main area has tabs for 'Overview', 'Assets' (highlighted with a blue underline), 'Jobs', and 'Manage'. On the left, there's a search bar, a 'Find assets' button, and a 'New asset +' button with a plus sign. Below it, there's a section for 'Asset types' with a note: 'After you create assets, they are organized by asset type.' On the right, there's a sidebar titled 'Data in this project' with a 'Drop data files here or browse for files to upload' area. The central content area shows '0 asset' and 'All assets' with a 'Start adding assets' callout and an illustration of a person interacting with a large plus sign icon.

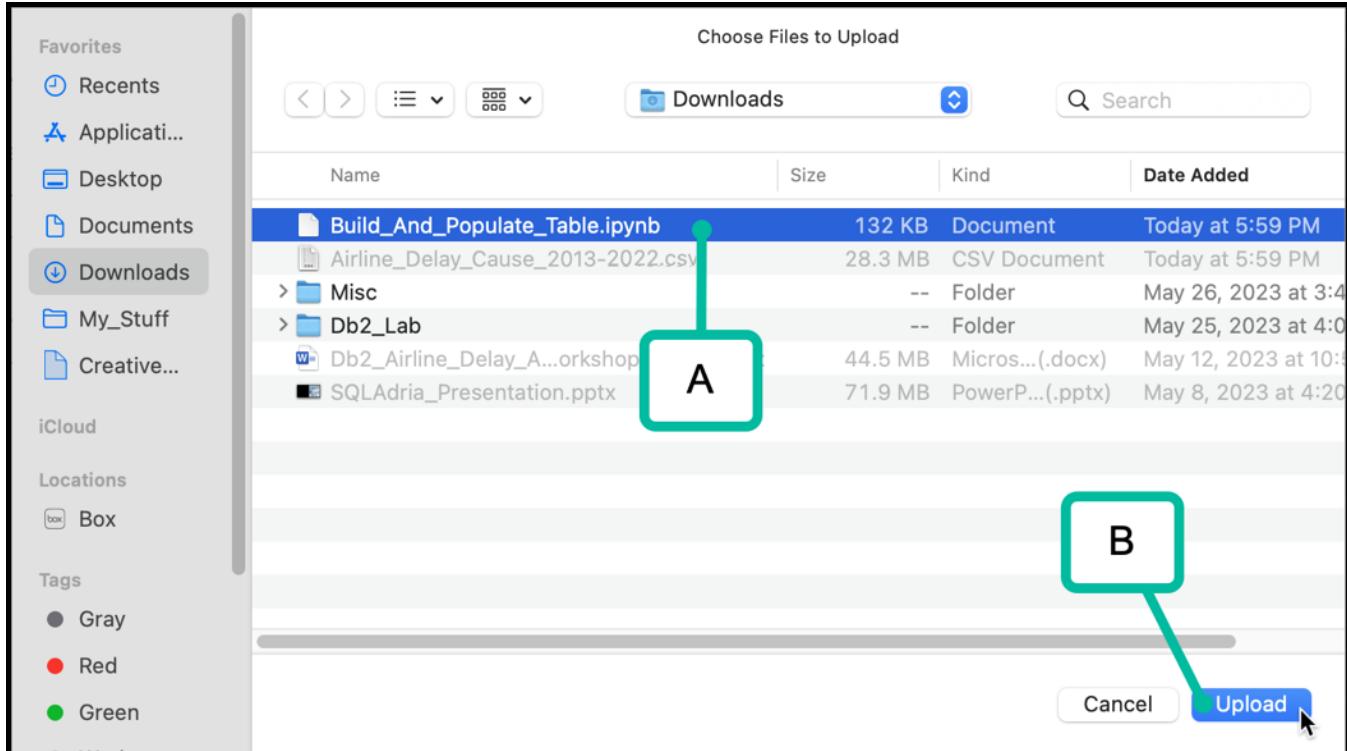
\_\_\_\_ 3. From the *New asset* popup window, [A] select the **</> Code editors** item in the **Tool type** menu shown on the left side of the screen. Then, [B] click on the **Jupyter notebook editor** tile. This will cause the *New asset* screen to be replaced with the *New notebook screen*.



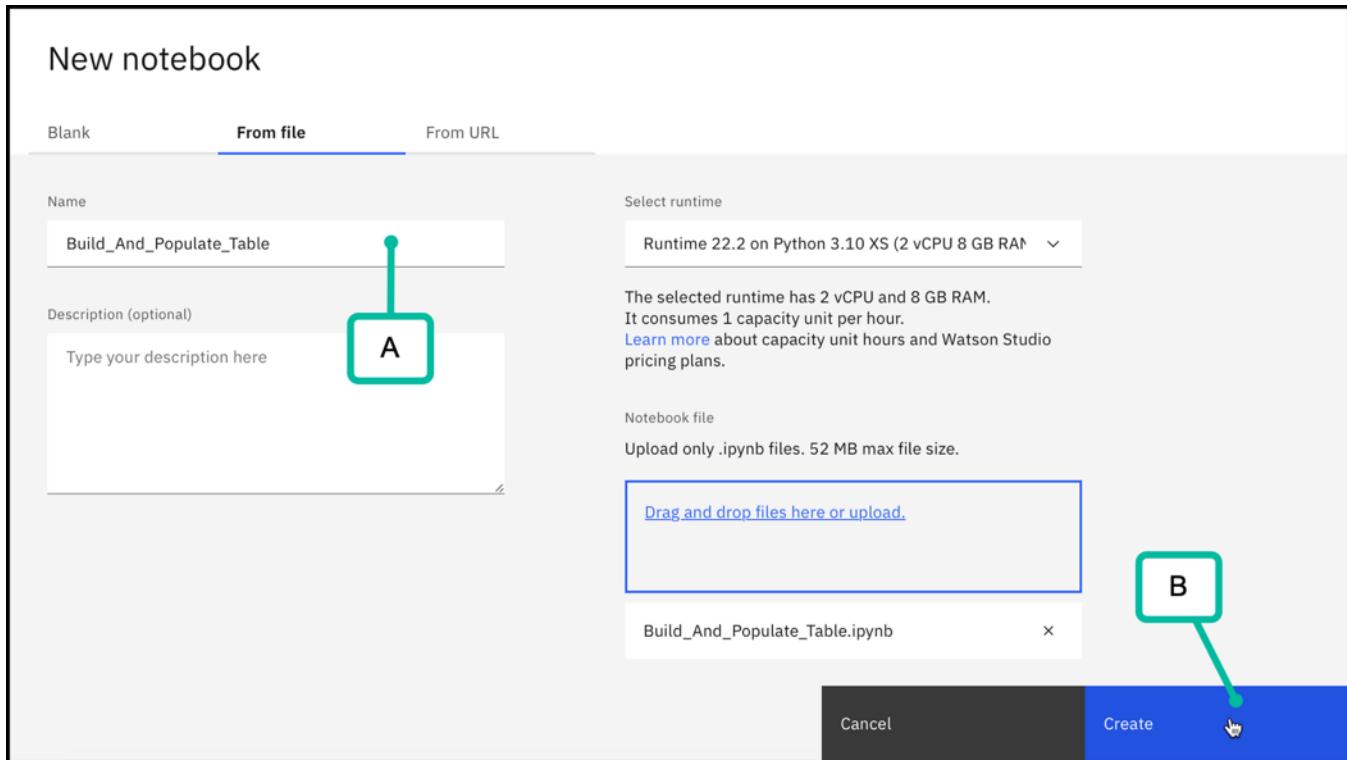
\_\_\_\_ 4. [A] Click on the **From file** tab of the *New notebook* screen. Then, [B] click on the **Drag and drop files here or upload** link to open a *Finder* window on Mac or a *File Explorer* window on Windows.



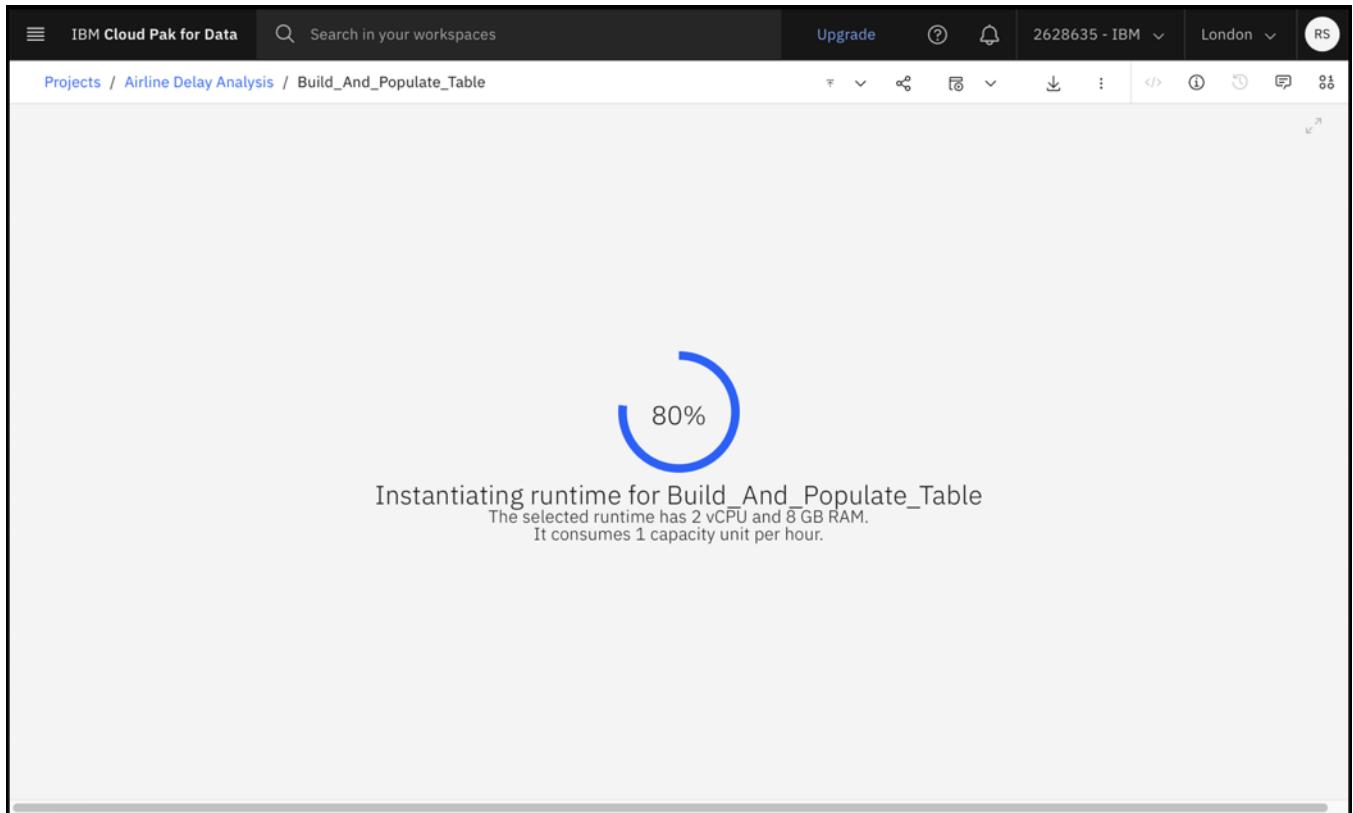
5. [A] Select the file named **Build\_And\_Populate\_Table.ipynb** and [B] click the **Upload** button.



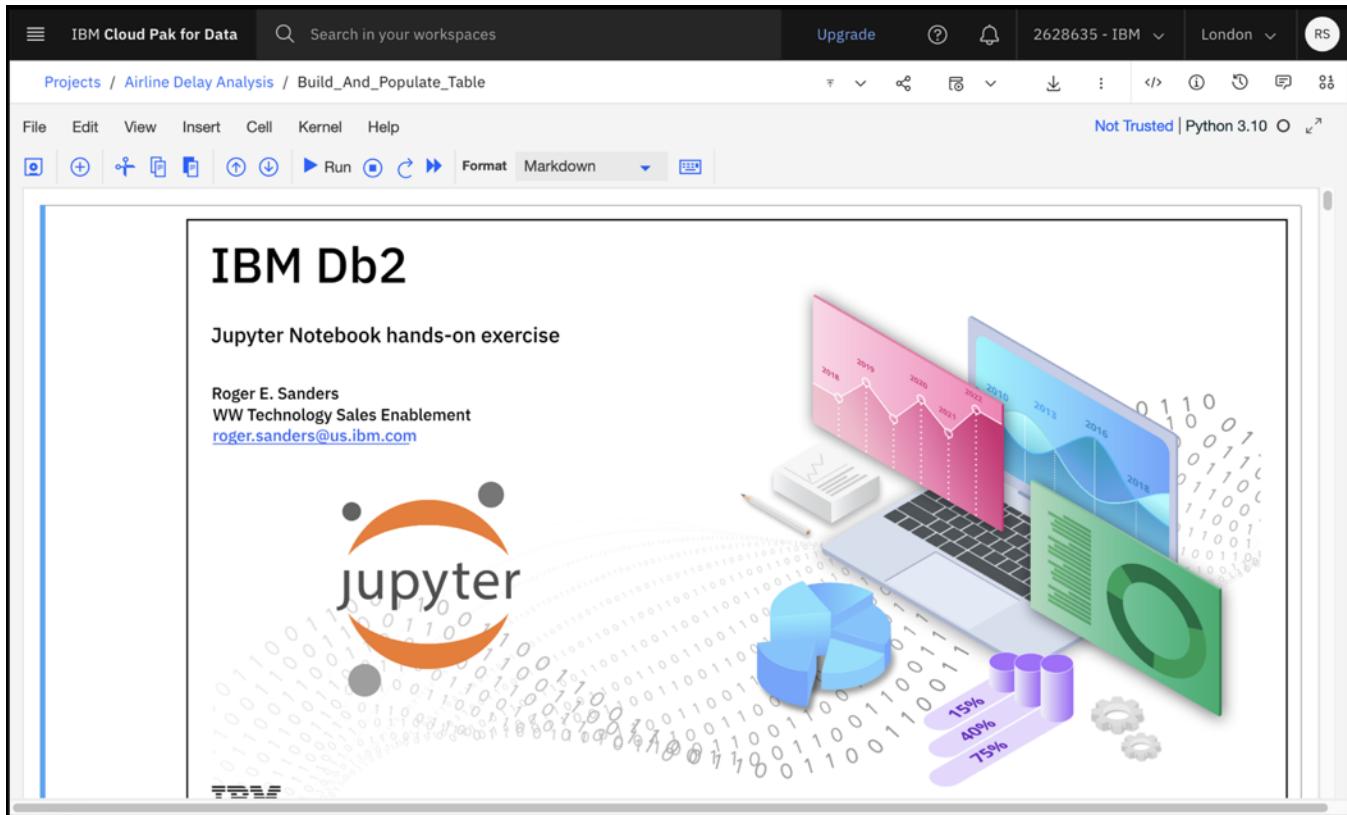
\_\_\_\_ 6. [A] Confirm that the **Name** field on the *New notebook* screen contains the name of the file just selected, minus the file extension (**Build\_And\_Populate\_Table**). Then, [B] click the **Create** button shown at the bottom of the screen.



\_\_\_\_ 7. The *New notebook* screen should be replaced with an *Instantiating runtime for [NotebookName]* screen that looks something like this:



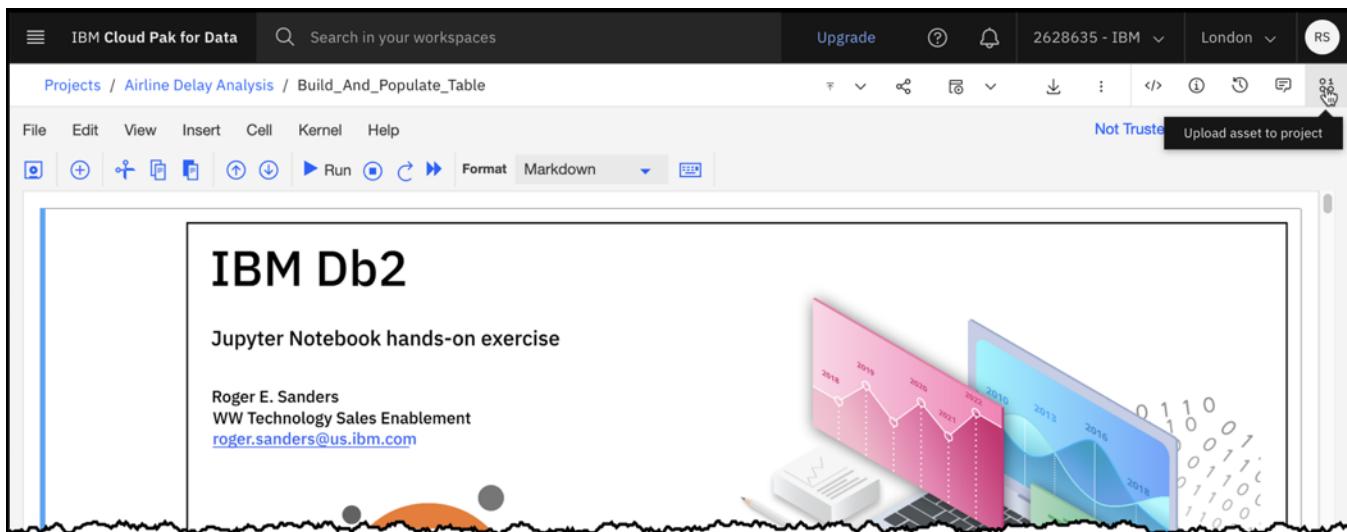
\_\_\_\_ 8. Once the runtime for the Jupyter Notebook that was loaded has been instantiated, the *Instantiating runtime for [NotebookName]* screen should be replaced with a screen that looks similar to this:



The Jupyter Notebook is now ready for editing (which we will do shortly).

## 4. Add a data file to the project.

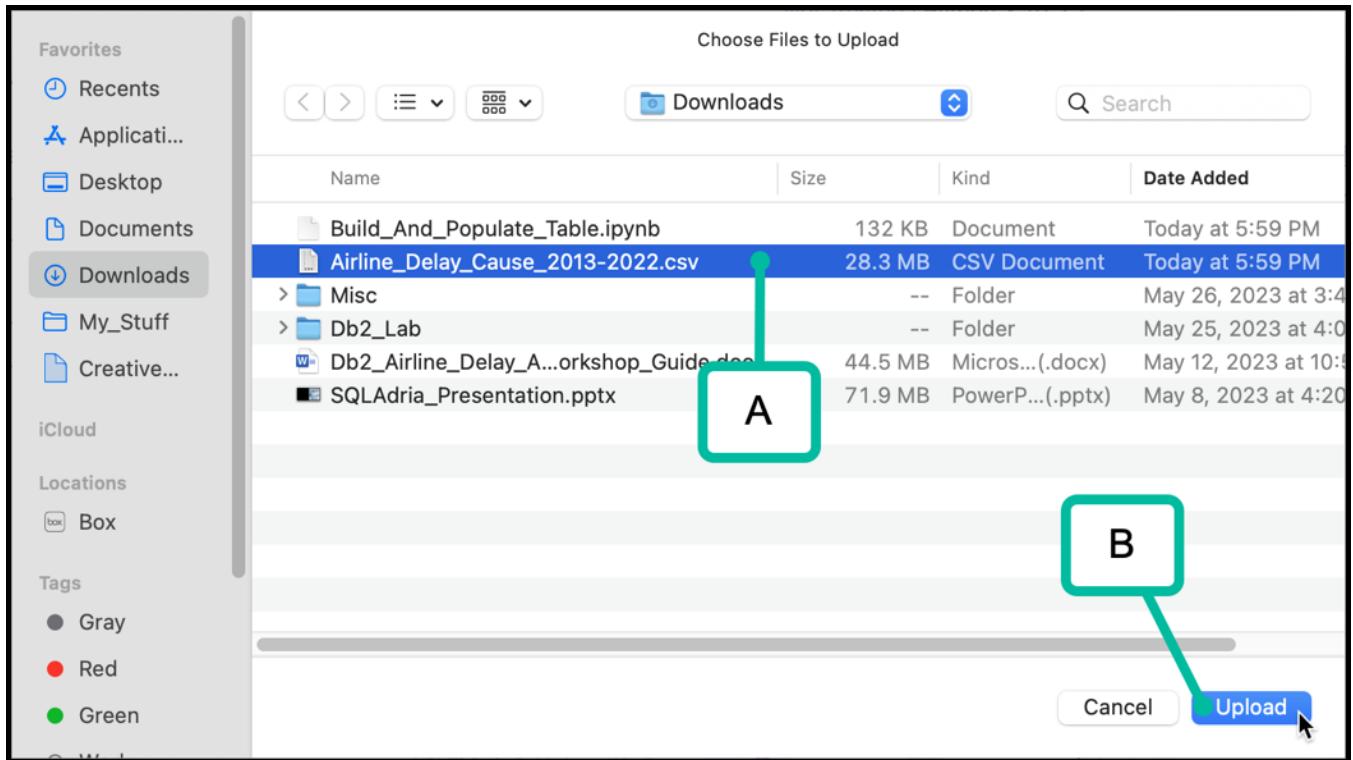
1. As the name implies, the **Build\_And\_Populate\_Table** Jupyter Notebook just instantiated is designed to create a new table in the Db2 on Cloud database created earlier, and then populate this table with data. But, before the notebook can be used to do that, a data source must be added to our project that the notebook can use. To do that, click on the **Upload asset to project** icon located in the top right-hand corner of the screen.



\_\_\_\_ 2. When the *Data in this project* slideout window appears, click on the **Drop data files here or browse for files to upload** link to open a *Finder* window on Mac or a *File Explorer* window on Windows.

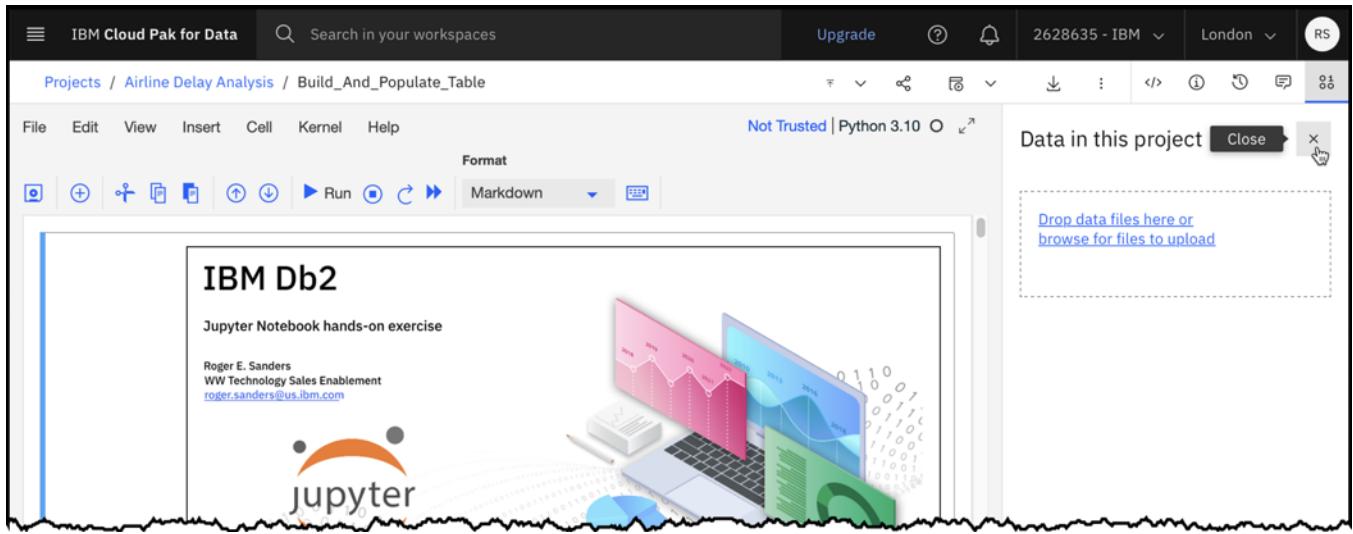
The screenshot shows the IBM Cloud Pak for Data interface. In the top right corner, there is a slideout window titled "Data in this project". Inside this window, there is a dashed box containing the text "Drop data files here or browse for files to upload". A cursor arrow is pointing towards this text. The main workspace shows a Jupyter Notebook titled "IBM Db2" with a subtitle "Jupyter Notebook hands-on exercise". The notebook contains text about Roger E. Sanders and the IBM Db2 database, along with various charts and binary code graphics. The top navigation bar includes tabs for File, Edit, View, Insert, Cell, Kernel, Help, and Format. The status bar at the bottom indicates "Not Trusted | Python 3.10".

\_\_\_\_ 3. [A] Select the file named **Airline\_Delay\_Cause\_2013-2022.csv** and [B] click the **Upload** button.



4. Once the appropriate data file has been selected, the *Data in this project* slideout window should look something like this:

5. After the data file has been successfully added to the project, close the *Data in this project* slideout window by clicking on the X (Close) button located in the top right-hand corner of the window.



**Congratulations!** You now have a Jupyter Notebook and a data asset in your **Airline Data Analysis** project. However, we need to make a couple of important alterations to the Jupyter Notebook before it can be run.

## 5. Prepare the Jupyter Notebook for execution.

1. Earlier, we learned that to work with an IBM Cloud resource from an application (such as a Jupyter Notebook), a service credential for that resource must exist. So, right after we created a Db2 on Cloud ‘Lite’ plan database, we created a service credential for that database (named **AirlineDelayDB**) and we copied the information associated with that service credential into a text file (named **AirlineDelayDB\_Credentials.txt**). Open the file that was saved earlier and locate the following information:

- [A] User name (“**username**”)
- [B] User password (“**password**”)
- [C] Database name (“**database**”)  
*NOTE: The database name is always **bludb** (BLUDB) for Db2 on Cloud databases.*
- [D] Host name (“**hostname**”)
- [E] Port number (“**port**”)

```

    ],
    "environment": {},
    "type": "cli"
},
"db2": {
    "authentication": {
        "method": "direct",
        "password": "N6ogXjFSeE8XOJF0",
        "username": "vxp49176"
    },
    "certificate": {
        "certificate_base64":
"LS0tLS1CRUdJTIBDRVJUSUZJQ0FURS0tLS0tCk1JSURIVENDQWdXZ0F3SUJBZ0lVT3dvMC9va09CUEN5RjFWefJxVGhKRW9ubDBVd0RRWUpLb1pJaHZTkFRRUwKQIFBd0hqRW
NNQm9HTFVRUF3d1RTVUpOSUVOCzIzVmJRVJ0ZEgavYITmxiekFIrncaeU1EQTRNRFF3TwpVMwpNaJphRncwek1EQTRNREI3TwpVM01qWmFnQjR4SERBYUJnTzCQU1NRT
BsQ1RTQkrRzkwVkcNCRVYUmhb2zU26ClpYTxDnZ0VpTUewRONTcUdTSWlzRFFFQkFRVUFBNECRHd8d2nRtBb0lCQVFEB0ZFNQ0SGdOeXZMUVlwR3gQKTBoamRXQnM4NV
BjTDNyRSjN1R3K2d1RudQSUxJU0VZV3o4Y1g1TG1XQk0rY1FnOG9VeSrQXJ3OEoxaxdRZQpySmllU21clf4WTM0c3BQeGRFVEZkWEhScnJhMGU2VmM4MW42TlJl02HSn1Q3hr
TG5GMUIFQW9hbHYwaDm2CnhDT0FvcXRwTfFrTzNpMTRGeU0yRDRIajkxckl4RGk4V9XMVpVdVhMNGwzZXXVLZUVCeTruZmhJv3kySVc3aUMKbGpMZ3RIN3hZTDVHbVpKOudsY
WtrSnj1cnpNREFQLzvUYnRIUlydEldotBRSVRFZHIESVFYUEZGRDBHYzloZao3M29JdpVZUJ3VC9uRHN80TJNC82SktzWpKN0lpdBFTN3Y2a2diUVhNDIBaUVJNxPQdUvpVz
NOY9GR0p0YCmY2a2jBZ01CQUFHaiV6QJNQjBHQTFRVZERnUvdCQIR2zZ2RU5MRjFvbWZnQ03MmxOcmMzSD12bURBUEJnTzUk1CQWY4RUJUQURBUUgvTUEwRwpDU3FHU0iIM0RRRUJDd1VBQTRJQkFRQtgydFvnUTZlaTZWHZndDj0dJdrbkpva1Y5JWNkaTNzbFVFWk
03MmxOcmMzSD12bURBUEJnTzUk1CQWY4RUJUQURBUUgvTUEwRwpDU3FHU0iIM0RRRUJDd1VBQTRJQkFRQtgydFvnUTZlaTZWHZndDj0dJdrbkpva1Y5JWNkaTNzbFVFWk
NDUytjCIVQZ3NmMnVBMDlxHiWTm1RkhjHZ1Vmp0VHRYTMk2NUM2WizRsnyxc3p1cu9zdFB5bkj4bln4cUs0dk0dtkVkjBWRUgcxe11znZBSmxkV3c4UEJ7ZGjtTk1HdGM4SzlwT
0050VdBQ1ZFRXVXVGdEhJKTXFBZnpYXlidUV0dwpoCw1pV2swTmVXNGk5ZEY4S2dTWUVAQWFodXVBSIRldXB2R2RPV1U0eEV4bm03aEVrbmZPV2ZITThDd08xNWFZCIRGQ
2s0Q0pDumR4Mlg5u284V3o1Z3MzncryRKFDQlJyZONYyeFDZnZrZTUDvVNHNkxFRHJHbmpWaXVSQkpZdW4Kt1RxWXROaVBHaHpuTHJrL0Fzam1LmzBxQmFLTmFyNuDQajhqlpN
b2RIZ04KLS0tLS1FTkQgQ0VSVEIGSUNBVEUtlS0tLQo=",
    "name": "dd14d0c-1b52-4f63-a606-53ecba28771d"
},
"composed": [
    "db2://vxg49176:N6ogXjFSeE8XOJF0@2d46b6b4-cbf6-40eb-bbce-6251e6ba0300.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:32328/bludb?
authSource=admin&replicaSet=replicaSet"
],
"database": "bludb",
"host_ros": [],
"hosts": [
{
    "hostname": "2d46b6b4-cbf6-40eb-bbce-6251e6ba0300.bs2io90l08kqb1od8lcg.databases.appdomain.cloud",
    "port": 32328
}
],
"jdbc_url": [
    "jdbc:db2://2d46b6b4-cbf6-40eb-bbce-6251e6ba0300.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:32328/
bludb:username=<userid>;password=<your_password>;sslConnection=true;"
]
}

```

2. In the Jupyter Notebook, [A] scroll down until you see the code section under the heading **Step 3a: Connect to a remote, Db2 on Cloud database**. Then, [B] click on this section to make it the current section. Next, copy the **hostname** value you found in the service credential file (**AirlineDelayDB\_Credentials.txt**) to the clipboard and paste it over the text “**replace-with-your-hostname**” found in this code section. (*Copy the value beside the **hostname** keyword in the service credential file and do NOT remove the quotation marks!*) Repeat this process for the **port** (“**replace-with-your-port-number**”), **username** (“**replace-with-your-userID**”), and **password** “**replace-with-your-password**” keywords. When you are finished, [C] click on the **Save Notebook** icon to make the changes made permanent. (NOTE: You can use the “hot keys” **command-C** on Mac or **Ctrl-C** on Windows to copy data to the clipboard; you can use the “hot keys” **command-V** on Mac or **Ctrl-V** on Windows to paste the copied data.)

**Step 3a: Connect to a remote, Db2 on Cloud database**

```
In [ ]: #-----#
# Initialize All User-Specific Connection Variables - Db2 on Cloud Database      #
#   IMPORTANT: UPDATE WITH VALUES FROM YOUR OWN ENVIRONMENT, AS PER LAB INSTRUCTIONS.  #
#-----#
# Define And Initialize The Appropriate Variables
dbName = "bludb"
hostName = "replace-with-your-hostname"
portNum = "replace-with-your-port-number"
userID = "replace-with-your-userID"
passWord = "replace-with-your-password"
secureComm = True           # Use SSL (Secure Sockets Layer) Communication

# Display A Status Message Indicating This Work Is Complete
print("\nUser-specific connection variable initialization work complete!\n")
```

**Step 3b: Connect to a local, on-premises Db2 database**

3. When finished, the code section under the heading **Step 3a: Connect to a remote, Db2 on Cloud database** should look something like this:

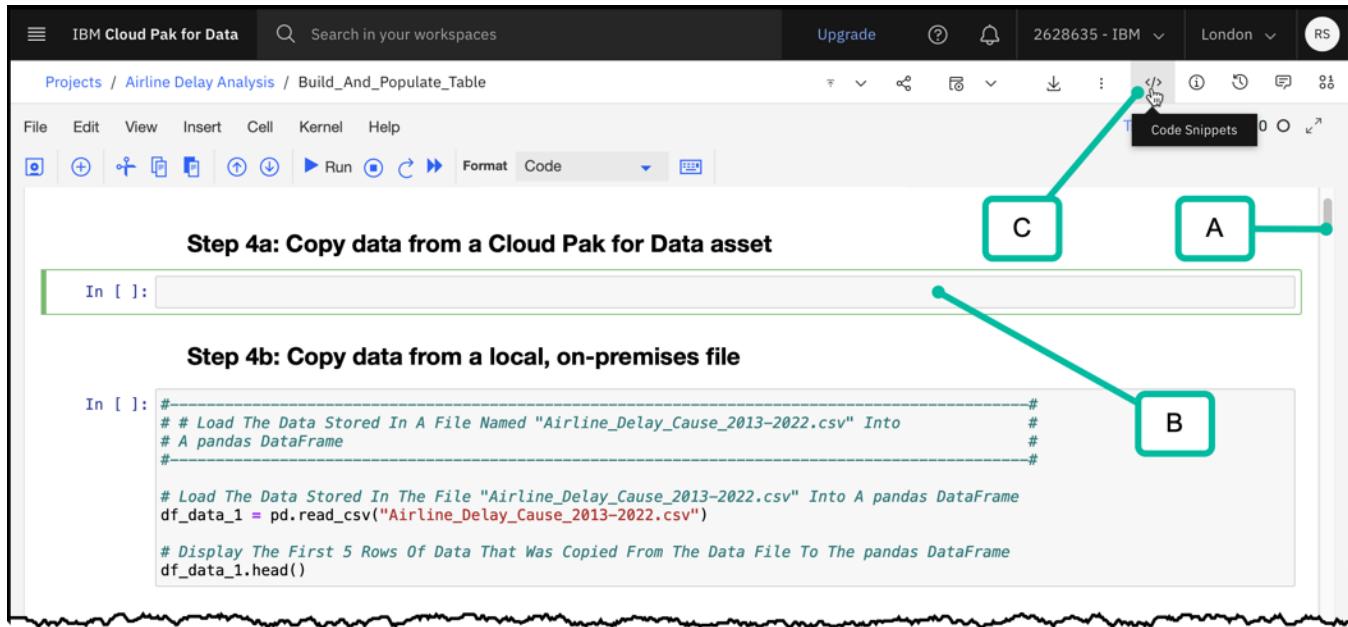
**Step 3a: Connect to a remote, Db2 on Cloud database**

```
In [ ]: #-----#
# Initialize All User-Specific Connection Variables - Db2 on Cloud Database      #
#   IMPORTANT: UPDATE WITH VALUES FROM YOUR OWN ENVIRONMENT, AS PER LAB INSTRUCTIONS.  #
#-----#
# Define And Initialize The Appropriate Variables
dbName = "bludb"
hostName = "2d46b6b4-cbf6-40eb-bbce-6251e6ba0300.bs2io90l08kqb1od8lcg.databases.appdomain.cloud"
portNum = "32328"
userID = "vxp49176"
passWord = "N6ogXjFSeE8X0JFo"
secureComm = True           # Use SSL (Secure Sockets Layer) Communication

# Display A Status Message Indicating This Work Is Complete
print("\nUser-specific connection variable initialization work complete!\n")
```

**Step 3b: Connect to a local, on-premises Db2 database**

4. [A] Scroll down until you see the code section under the heading **Step 4a: Copy data from a Cloud Pak for Data asset**. Next, [B] click on this section to make it the current section. Finally, [C] click on the </> (Code Snippets) icon located near the top right corner of the screen.



5. When the *Code Snippets* slideout window appears on the right-hand side of the screen, locate the **Read data** tile (under the heading **Data Ingestion**) and click on it.

**Step 4a: Copy data from a Cloud Pak for Data asset**

**Step 4b: Copy data from a local, on-premises file**

```
In [ ]: # Load The Data Stored In A File Named "Airline_Delay_Cause_2013-2022.csv" Into A pandas DataFrame  
# Load The Data Stored In The File "Airline_Delay_Cause_2013-2022.csv" Into A pandas DataFrame  
df_data_1 = pd.read_csv("Airline_Delay_Cause_2013-2022.csv")  
# Display The First 5 Rows Of Data That Was Copied From The Data File To The pandas DataFrame  
df_data_1.head()
```

**Step 5: Copy data stored in a pandas DataFrame into the AIRLINE\_DELAY\_CAUSE\_TOTAL table (using a prepared INSERT statement with parameter markers)**

**Overview:**

When a table is first created, it is nothing more than a definition of how a set of data values are to be stored.

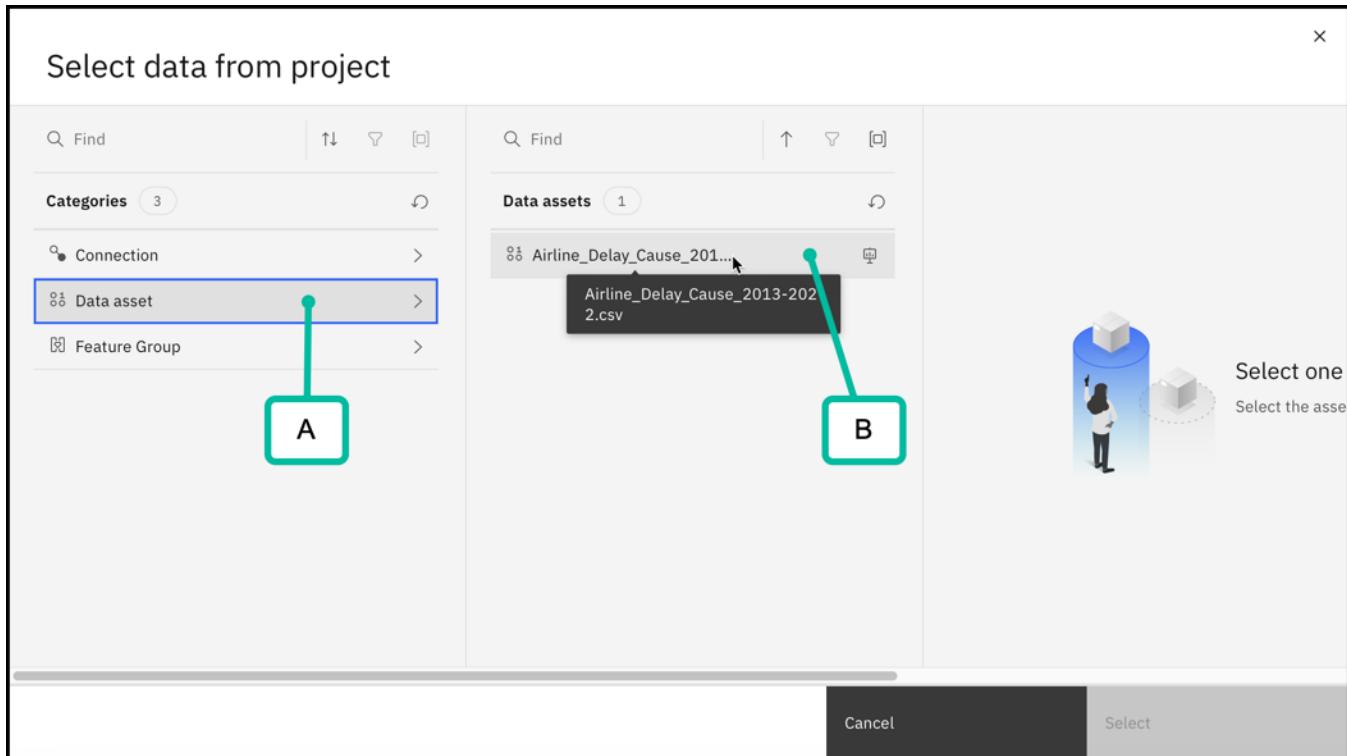
6. When the *Code Snippets* slideout window is replaced with the *Read data* window, click on the **Select data from project** button. This will cause the *Select data from project* popup window to open.

The screenshot shows a Jupyter Notebook interface within the IBM Cloud Pak for Data environment. The top navigation bar includes 'IBM Cloud Pak for Data', a search bar, 'Upgrade', '2628635 - IBM', 'London', and a user icon. The main area displays a notebook titled 'Build\_And\_Populate\_Table' under the 'Airline Delay Analysis' project. The code cell contains the following Python code:

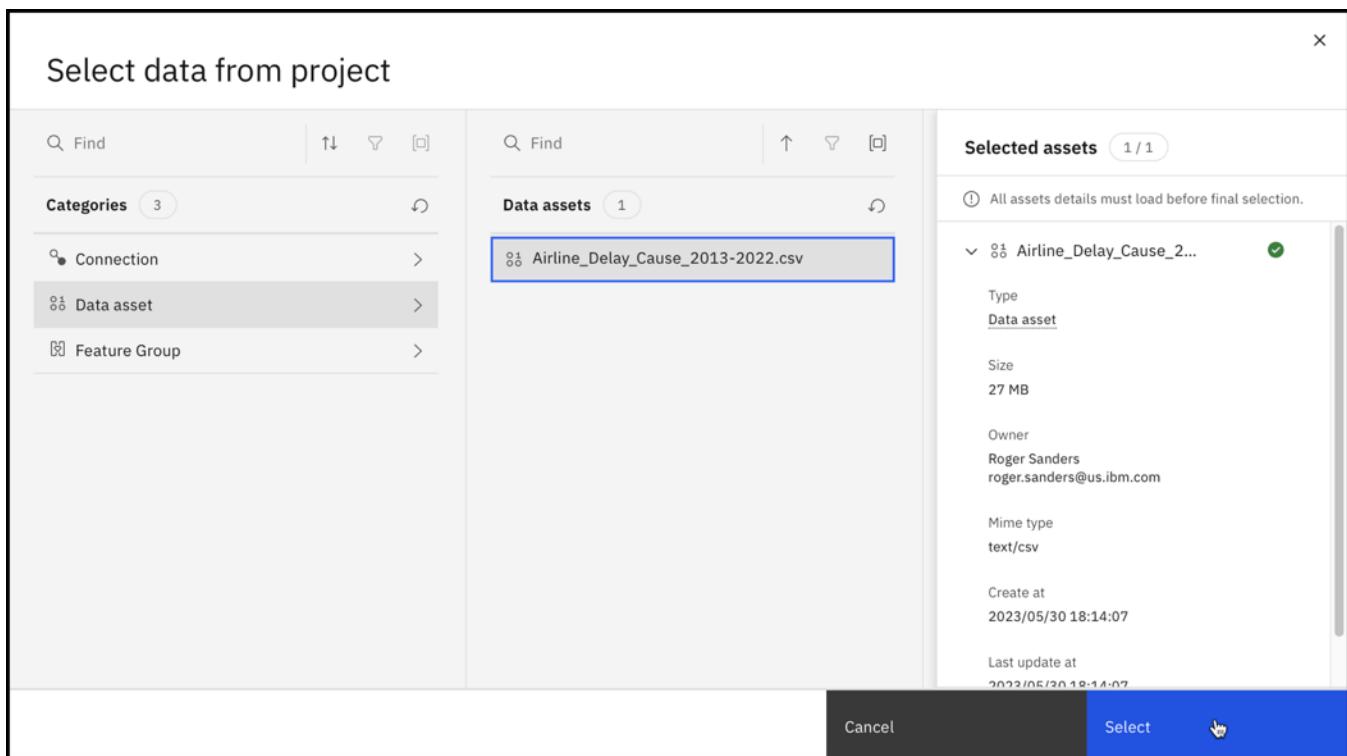
```
In [ ]: # Load The Data Stored In A File Named "Airline_Delay_Cause_2013-2022.csv" Into A pandas DataFrame
# Load The Data Stored In The File "Airline_Delay_Cause_2013-2022.csv" Into A pandas DataFrame
df_data_1 = pd.read_csv("Airline_Delay_Cause_2013-2022.csv")
# Display The First 5 Rows Of Data That Was Copied From The Data File To The pandas DataFrame
df_data_1.head()
```

Below the code cell, a section titled 'Step 4a: Copy data from a Cloud Pak for Data asset' is visible. To the right, a sidebar titled 'Read data' provides instructions to 'Generate a code snippet to load data from a data asset or connection into your notebook' and includes a 'Select data from project' button.

7. When the **Select data from project** popup window appears, **[A]** locate the **Data asset** item in the **Categories** list and click on it. This will cause a list of all **Data assets** that have been defined for the **Airline Delay Analysis** project to be displayed. When this list appears, **[B]** locate the **Airline\_Delay\_Cause\_2013-2022.csv** item and click on it to select it. (This should be the only item in the list.)



8. Review the details provided about the data asset selected (shown under the **Selected assets** heading) and confirm the correct data asset was chosen. Then, click the **Select** button located in the bottom right corner of the window. This will close the popup window and return control to the *Read data* slideout window.



\_\_\_\_ 9. The data asset just selected should be displayed under the **Selected data** heading and the **Load as** dropdown/field should contain the value **pandas DataFrame**. (If that's not the case, click on the **Load as** dropdown and select **pandas DataFrame** from the list provided). Click the **Insert code to cell** button shown at the bottom of the window. This will cause the code that is needed to access the asset chosen to be generated in the code section under the heading **Step 4a: Copy data from a Cloud Pak for Data asset**.

The screenshot shows the IBM Cloud Pak for Data interface. At the top, there is a navigation bar with 'IBM Cloud Pak for Data', a search bar, and various system status indicators. Below the navigation bar is a toolbar with file operations like 'File', 'Edit', 'View', 'Insert', 'Cell', 'Kernel', and 'Help'. To the right of the toolbar is a status bar showing 'Trusted | Python 3.10' and other details. The main workspace contains a Jupyter Notebook cell labeled 'In [ ]:' which displays the following code:

```
In [ ]: # Load The Data Stored In A File Named "Airline_Delay_Cause_2013-2022.csv" Into  
# A pandas DataFrame  
#  
# Load The Data Stored In The File "Airline_Delay_Cause_2013-2022.csv" Into A pandas Da  
df_data_1 = pd.read_csv("Airline_Delay_Cause_2013-2022.csv")  
# Display The First 5 Rows Of Data That Was Copied From The Data File To The pandas Da  
df_data_1.head()
```

Below the code cell, there is a section titled 'Step 4a: Copy data from a Cloud Pak for Data asset' and another titled 'Step 4b: Copy data from a local, on-premises file'. To the right of the workspace, a slideout window titled 'Read data' is open. It shows the selected data asset 'Airline\_Delay\_Cause\_2013-2022.csv' and the 'Load as' dropdown set to 'pandas DataFrame'. At the bottom right of the slideout window is a blue button labeled 'Insert code to cell' with a cursor icon pointing to it.

\_\_\_\_ 10. Once the appropriate code has been generated, close the *Read data* slideout window by clicking on the X (Close) button located in the top right-hand corner of the window.

The screenshot shows a Jupyter Notebook interface within the IBM Cloud Pak for Data environment. The notebook is titled "Build\_And\_Populate\_Table". The code cell contains Python code for reading a CSV file from an IBM Cloud Object Storage bucket. The code includes imports for os, types, pandas, botocore.client, Config, and ibm\_boto3. It defines a class with a \_\_iter\_\_ method and uses the pandas.read\_csv function to read the file. A note in the code indicates it's reading from a file in IBM Cloud Object Storage.

```

In [ ]:
import os, types
import pandas as pd
from botocore.client import Config
import ibm_boto3

def __iter__(self): return 0

# @hidden_cell
# The following code accesses a file in your IBM Cloud Object Storage. It includes you
# You might want to remove those credentials before you share the notebook.
cos_client = ibm_boto3.client(service_name='s3',
    ibm_api_key_id='eAo5nLJDozlDismMGk4WnA8cIvWD7x06uVFvCu_MxSC',
    ibm_auth_endpoint="https://iam.cloud.ibm.com/oidc/token",
    config=Config(signature_version='oauth'),
    endpoint_url='https://s3.private.eu-gb.cloud-object-storage.appdomain.cloud')

bucket = 'airlinedelayanalysis-donotdelete-pr-d5spreb4lhyl78'
object_key = 'Airline_Delay_Cause_2013-2022.csv'

body = cos_client.get_object(Bucket=bucket,Key=object_key)['Body']
# add missing __iter__ method, so pandas accepts body as file-like object
if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType( __iter__, body )

df_data_2 = pd.read_csv(body)
df_data_2.head()

```

\_\_\_\_ 11. Edit the code that was generated and stored in the code section under the heading **Step 4a: Copy data from a Cloud Pak for Data asset** as follows:

\_\_\_\_ 11a. [A] Remove the lines “**import os, types**“ and “**import pandas as pd**“. (These modules are imported in the first code cell in the Jupyter Notebook so there’s no need to import them again here.)

\_\_\_\_ 11b. [B] Make sure the two references to the variable **df\_data\_x** are set to **df\_data\_1**. (If you look at the screenshot shown in the previous step, you’ll notice that a variable named **df\_data\_2** was generated. The variable name **df\_data\_1** is referenced in other code cells, so it must be used here.)

When you are finished, [C] click on the **Save Notebook** icon located in the top right-hand corner of the screen to make the changes made permanent.

The screenshot shows a Jupyter Notebook interface in IBM Cloud Pak for Data. The title bar indicates the workspace is 'Airline Delay Analysis / Build\_And\_Populate\_Table'. The toolbar includes 'File', 'Edit', 'View', 'Insert' (with a highlighted 'Cell' icon), 'Kernel', and 'Help'. The status bar shows 'Trusted | Python 3.10'.

**Step 4a: Copy data from a Cloud Pak for Data asset**

```
In [ ]:
from botocore.client import Config
import ibm_boto3

def __iter__(self): return 0

# @hidden_cell
# The following code accesses a file in your IBM Cloud Object Storage. It includes your credentials.
# You might want to remove those credentials before you share the notebook.
cos_client = ibm_boto3.client(service_name='s3',
    ibm_api_key_id='eAo5nLJDozLDlismMGk4WnA8c1vWD7xQ6uVFvCu_MxSC',
    ibm_auth_endpoint="https://iam.cloud.ibm.com/oidc/token",
    config=Config(signature_version='oauth'),
    endpoint_url='https://s3.private.eu-gb.cloud-object-storage.appdomain.cloud')

bucket = 'airlinedelayanalysis-donotdelete-pr-d5speb4lhy1y78'
object_key = 'Airline_Delay_Cause_2013-2022.csv'

body = cos_client.get_object(Bucket=bucket,Key=object_key)['Body']
# add missing __iter__ method, so pandas accepts body as file-like object
if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType( __iter__, body )

df_data_1 = pd.read_csv(body)
df_data_1.head()
```

**Step 4b: Copy data from a local, on-premises file**

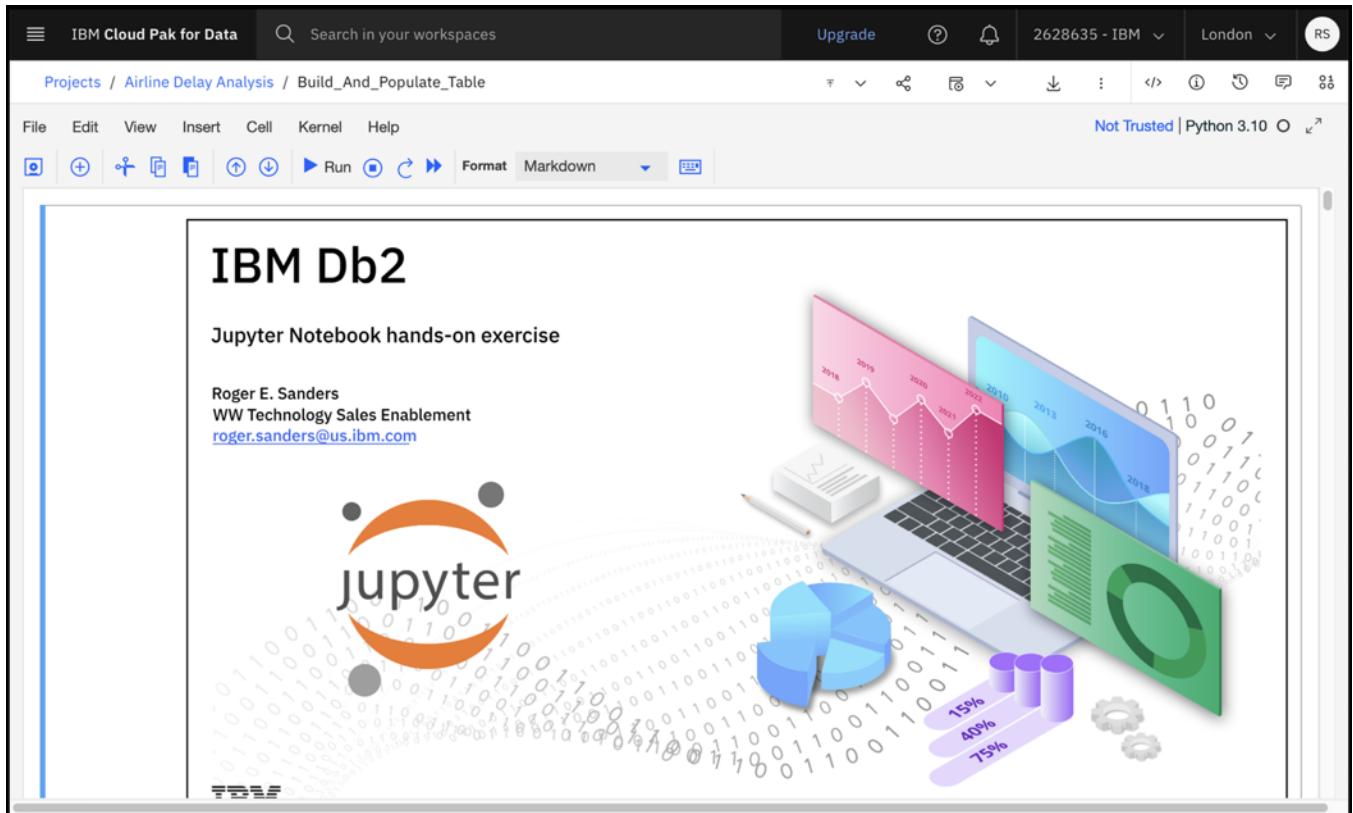
The code cell contains annotations: 'C' points to the 'Cell' icon in the toolbar; 'A' points to the first line of code; 'B' points to the line 'df\_data\_1.head()'.

**Congratulations!** You are now ready to work your way through the Jupyter Notebook.

## 6. Run the Jupyter Notebook

1. Scroll back to the beginning of the Jupyter Notebook and work your way through each cell, one cell at a time. The notebook is clearly documented and easy to understand.

**IMPORTANT: Do NOT execute Section 3, Step 1: Remove the AIRLINE\_DELAY\_CAUSE\_TOTAL table created earlier! This table will be needed for the next part of this workshop. Although this step is not meant to be used, it is included in the notebook to show you how to drop a Db2 on Cloud database table when it is no longer needed.**



### Workshop Tip

If your Jupyter Notebook appears to be non-responsive, you can “reboot” it and start over from the beginning by selecting **Kernel** from the menu bar shown at the top of the screen and choosing the option **Restart and Clear Output** from the dropdown menu presented.

This will cause a confirmation popup window to be displayed, and you must confirm the operation before the Jupyter Notebook kernel can be restarted.

\_\_\_\_ 2. When you are finished, log out of IBM Cloud Pak for Data as a Service by [A] clicking on your user icon located in the top right-hand corner of the screen and then [B] clicking on the **Log out** item in the menu presented.

The screenshot shows the IBM Cloud Pak for Data Jupyter Notebook interface. The top navigation bar includes 'IBM Cloud Pak for Data', a search bar, 'Upgrade', user profile '2628635 - IBM', location 'London', and a 'Log out' button. The main content area displays a Jupyter notebook titled 'IBM Db2' with the subtitle 'Jupyter Notebook hands-on exercise'. It features a bio for 'Roger E. Sanders' and his email 'roger.sanders@us.ibm.com'. To the right of the notebook is a decorative graphic of a server tower with a 3D bar chart and binary code. Callouts labeled 'A' and 'B' point to the 'Log out' button and the server/binary graphic respectively.

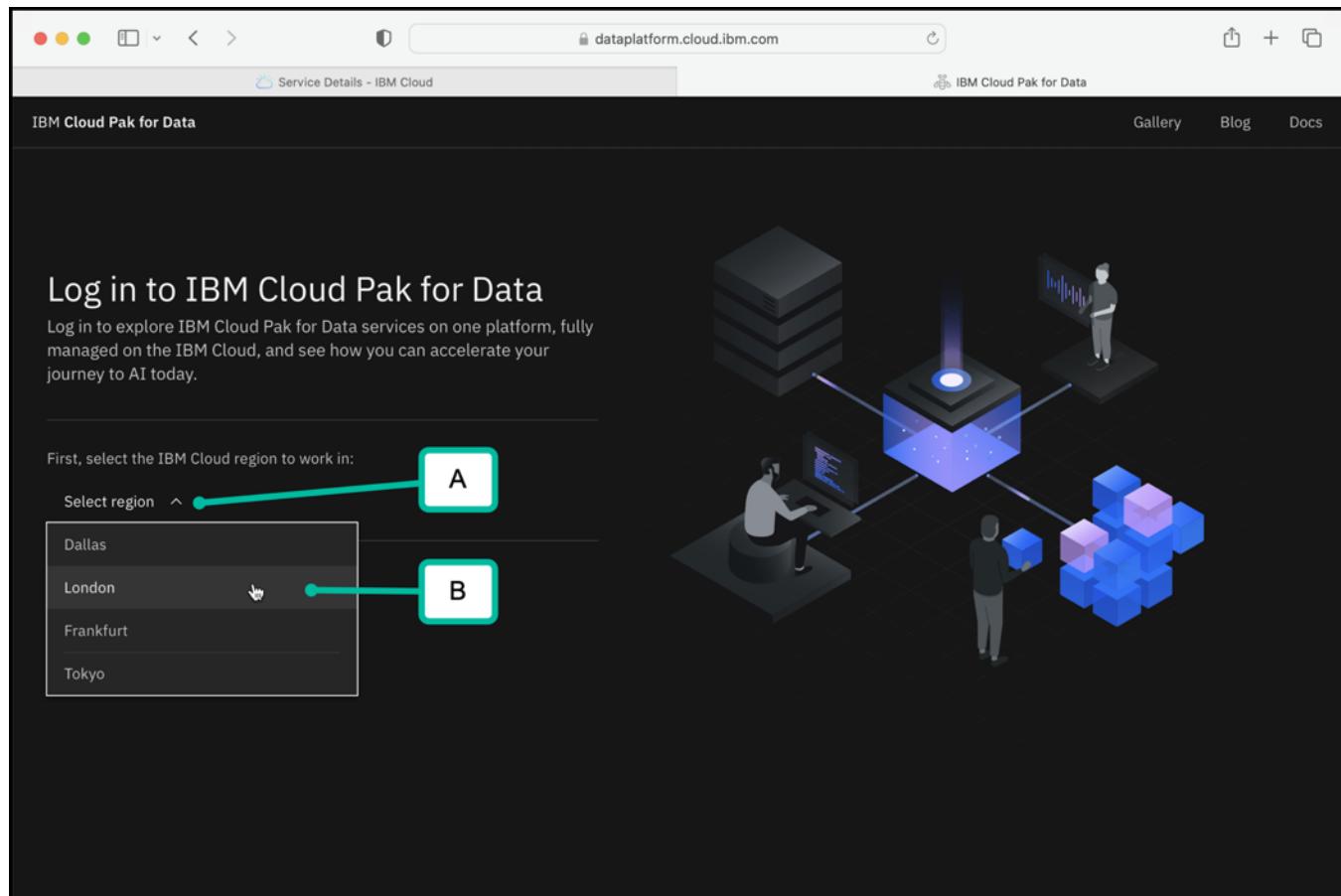
**Congratulations!** You have just used IBM Watson Studio (in IBM Cloud Pak for Data as a Service) and Jupyter Notebook to work with a Db2 on Cloud database.

## V. Analyzing the data

In this section of the workshop, you will use another Jupyter Notebook to analyze the airline flight delay data that was loaded into the Db2 on Cloud database. You will also use the Db2 Magic Commands to establish a database connection and retrieve data from the AIRLINE\_DELAY\_CAUSE and AIRLINE\_DELAY\_CAUSE\_TOTAL tables.

### 1. Log in to your IBM Cloud Pak for Data as a Service account.

\_\_\_\_ 1. Open a web browser and go to the *Log in to BM Cloud Pak for Data* web site (<https://dataplatform.cloud.ibm.com/login>). Once there, [A] click on the **Select region** button to display a dropdown list of locations where Cloud Pak for Data as a Service can be deployed. Then, [B] choose the location where your IBM Cloud Pak as a Service account is to be provisioned. (In this example, the **London** location was selected.)



\_\_\_\_ 2. When the log in fields appear, make sure the **Log in with** field contains the value **IDBId**. Then, enter the email address you used to create your IBMid in the **IBMid** field (located beside the **Log in with** field).

**IMPORTANT: Please use the same IBMid that was used earlier to create an IBM Cloud account and provision a Db2 on Cloud Lite plan database.** If prompted, provide the password associated with your IBMid and then click the **Continue** button. (You may be asked to log in to IBM as well.)

IBM Cloud Pak for Data

Gallery Blog Docs

## Log in to IBM Cloud Pak for Data

Log in to explore IBM Cloud Pak for Data services on one platform, fully managed on the IBM Cloud, and see how you can accelerate your journey to AI today.

Work with resources in this IBM Cloud region

London

Log in with IBMID Forgot IBMID?

IBMid

Continue 

Remember IBMID

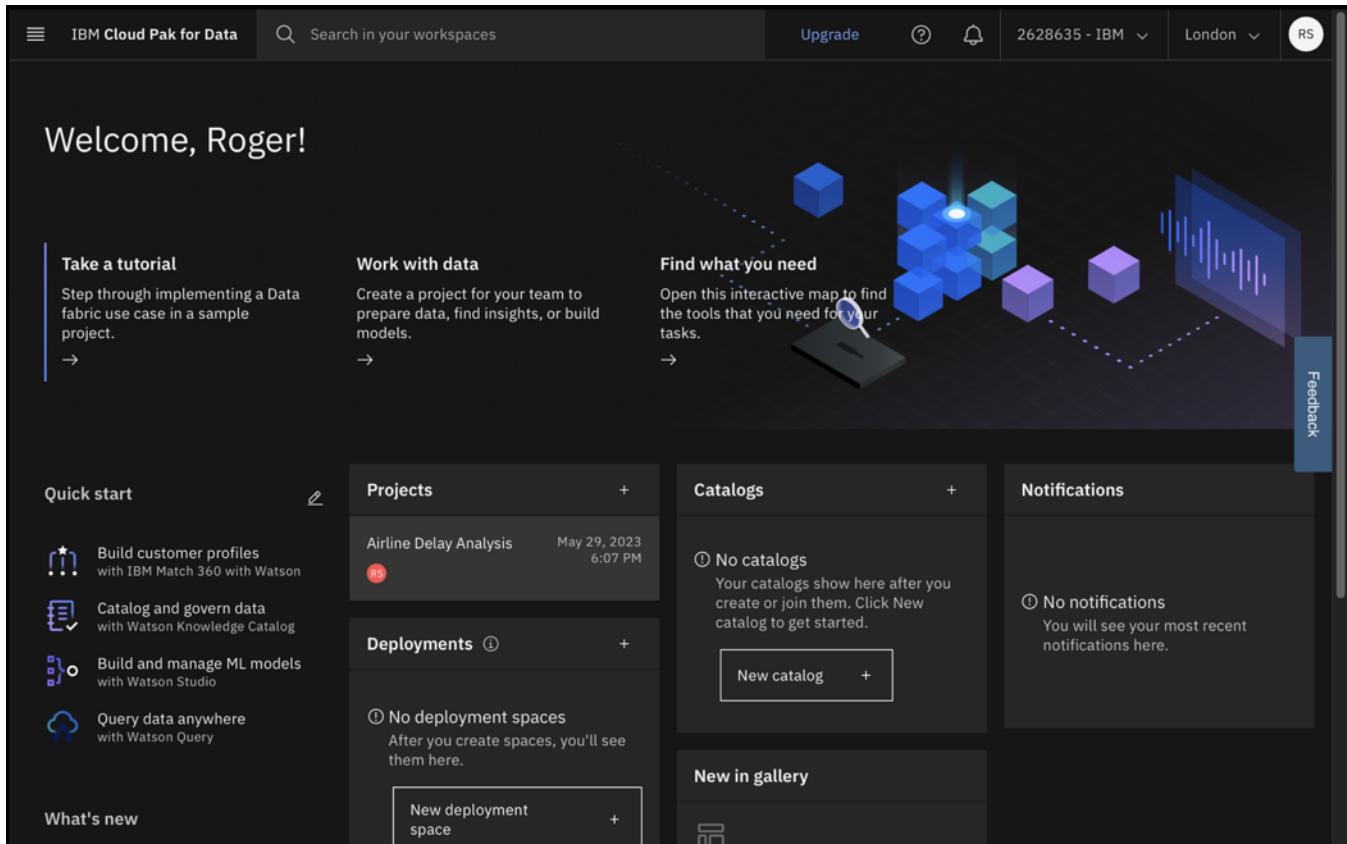
Enable Cloud Foundry access 

Need an account? [Sign up and try for free](#)

Need help?



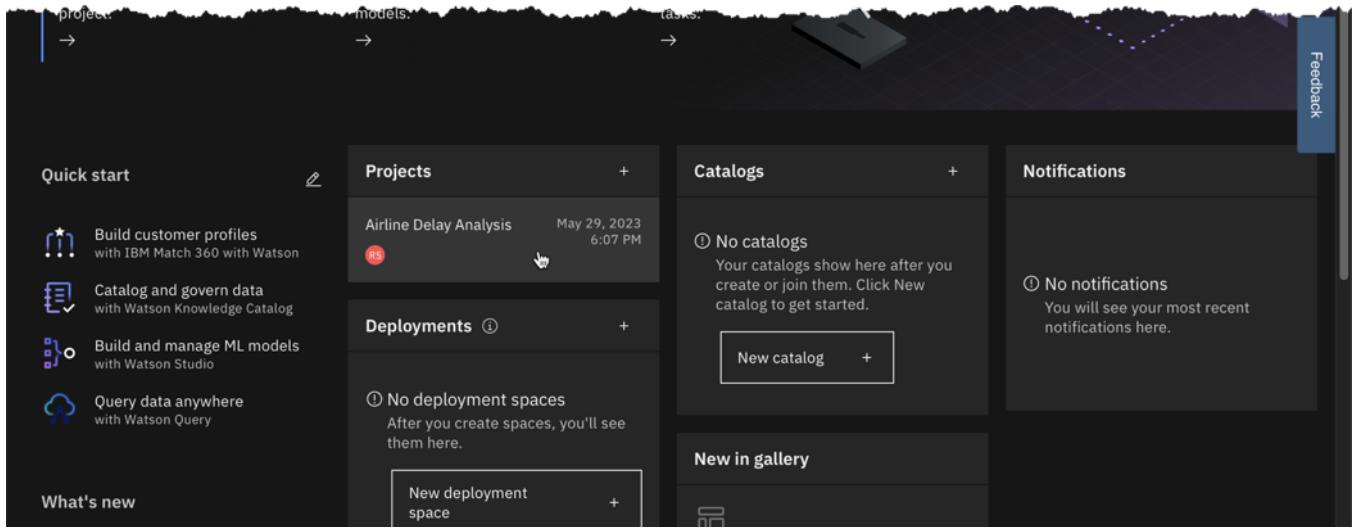
\_\_\_\_ 3. Once you have successfully logged in, you should be presented with an IBM Cloud Pak for Data as a Service *Welcome* screen that looks something like this:



The screenshot shows the IBM Cloud Pak for Data welcome screen. At the top, there's a navigation bar with 'IBM Cloud Pak for Data', a search bar, 'Upgrade' button, and account information ('2628635 - IBM' and 'London'). Below the header is a 'Welcome, Roger!' message. On the left, there are three main sections: 'Take a tutorial' (with a link to a sample project), 'Work with data' (with a link to a project creation page), and 'Find what you need' (with a link to an interactive map). The central area features a 'Quick start' section with four items: 'Build customer profiles with IBM Match 360 with Watson', 'Catalog and govern data with Watson Knowledge Catalog', 'Build and manage ML models with Watson Studio', and 'Query data anywhere with Watson Query'. To the right, there are three cards: 'Projects' (listing 'Airline Delay Analysis'), 'Catalogs' (showing 'No catalogs' and a 'New catalog' button), and 'Notifications' (showing 'No notifications' and a 'New in gallery' button). A vertical 'Feedback' button is located on the far right.

## 2. Add a new Jupyter Notebook to the project created earlier.

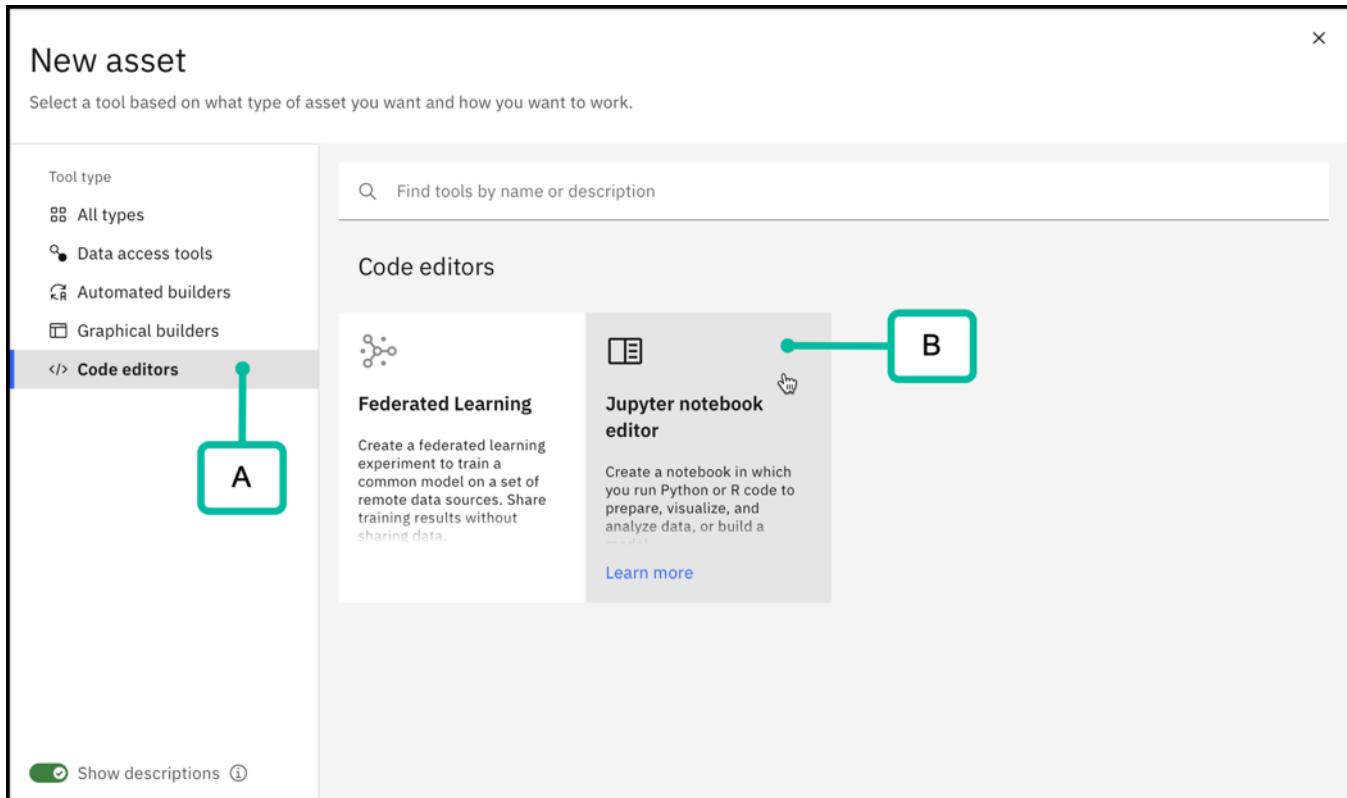
1. When the IBM Cloud Pak for Data as a Service *Welcome* screen appears, locate the project in the **Projects** box that was created earlier (in this example, the project was named “*Airline Delay Analysis*”) and click on it. This should cause the *Welcome* screen to be replaced with an **Assets** screen that shows the assets that have been created for the project selected. (If an *Overview* screen is presented instead, click on the **Assets** tab at the top of the screen to switch to the **Assets** screen.)



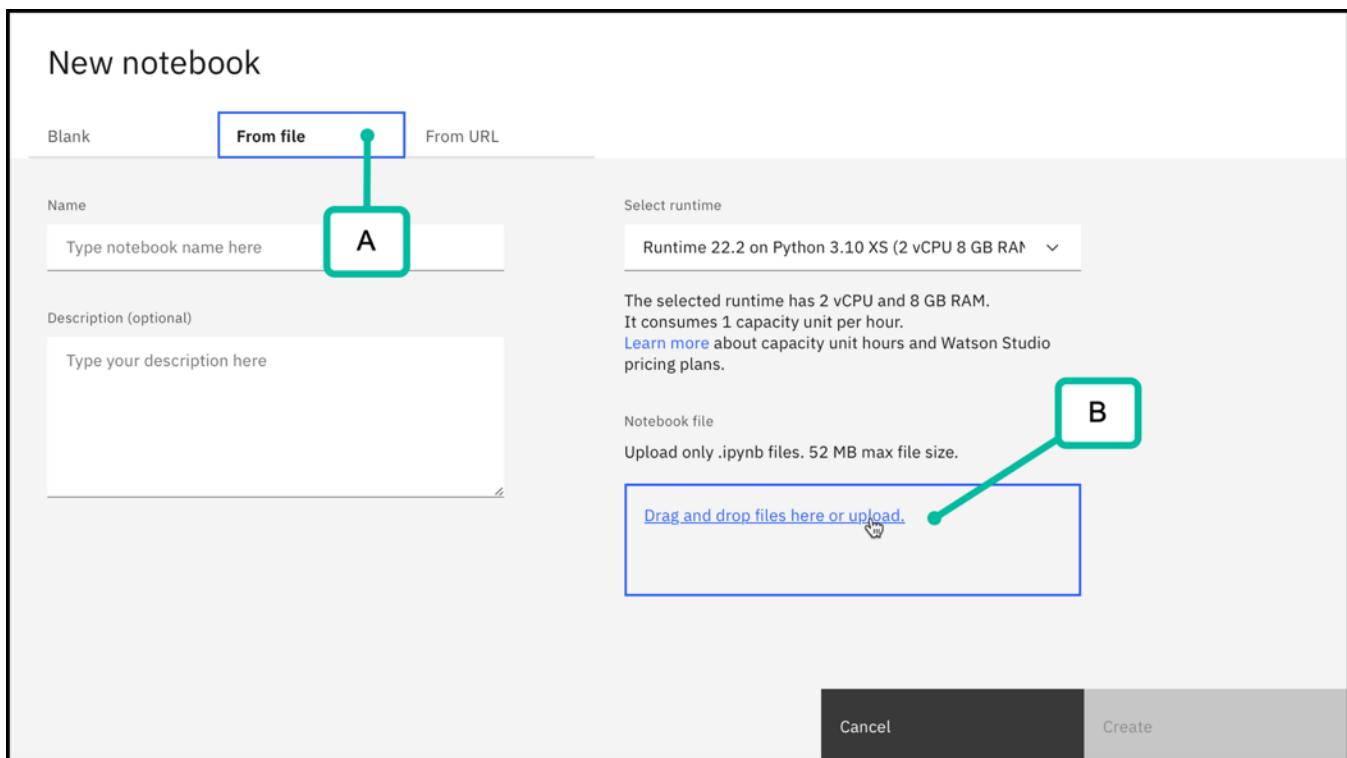
2. When the Assets screen is displayed, click on the **New asset** button.

The screenshot shows the 'Assets' screen in the 'Airline Delay Analysis' project. The 'Assets' tab is selected. On the right, there's a 'New asset' button with a '+' icon. The left sidebar shows '2 assets' and 'Asset types' (Data and Notebooks). The main area shows a table of 'All assets' with two entries: 'Build\_And\_Populate\_Table' (Notebook) and 'Airline\_Delay\_Cause\_2013-20...' (CSV).

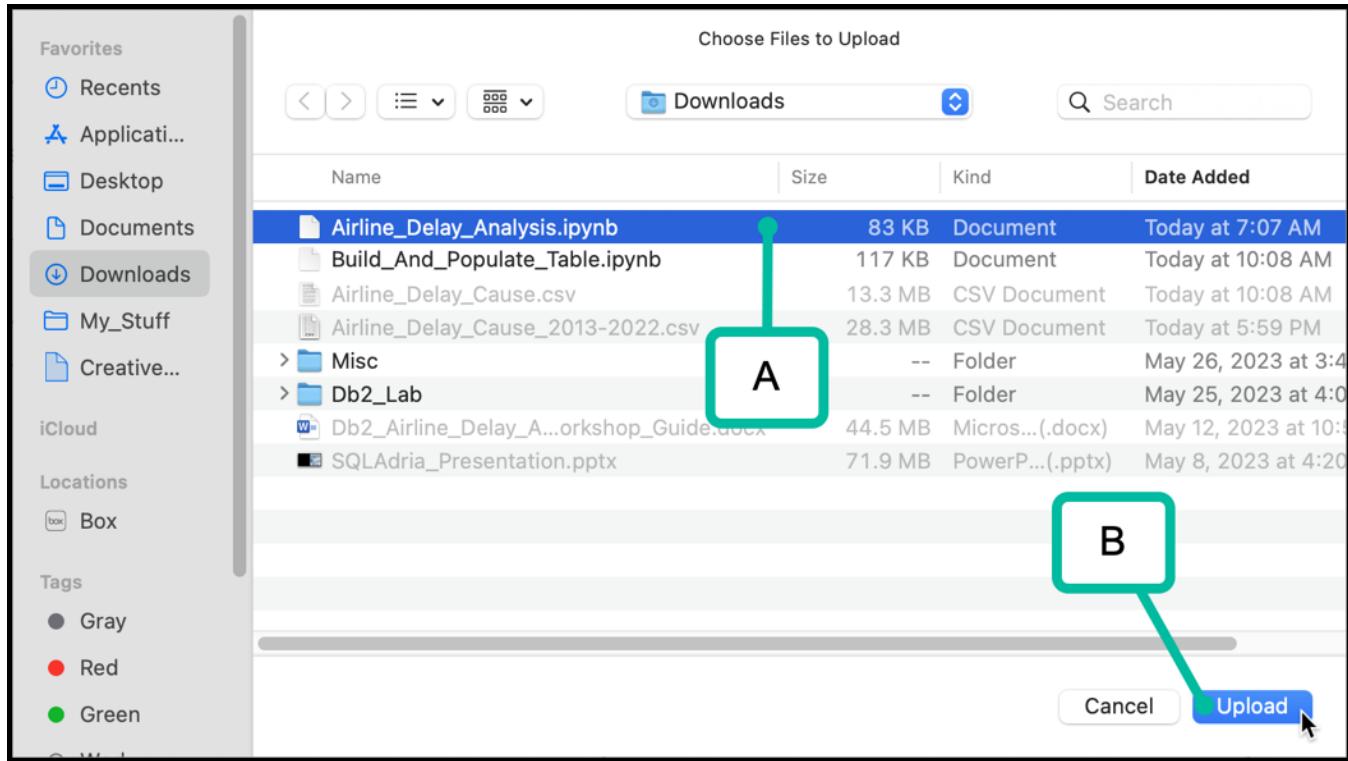
3. From the *New asset* popup window, [A] select the **</> Code editors** item in the **Tool type** menu shown on the left side of the screen. Then, [B] click on the **Jupyter notebook editor** tile. This will cause the *New asset* screen to be replaced with the *New notebook* screen.



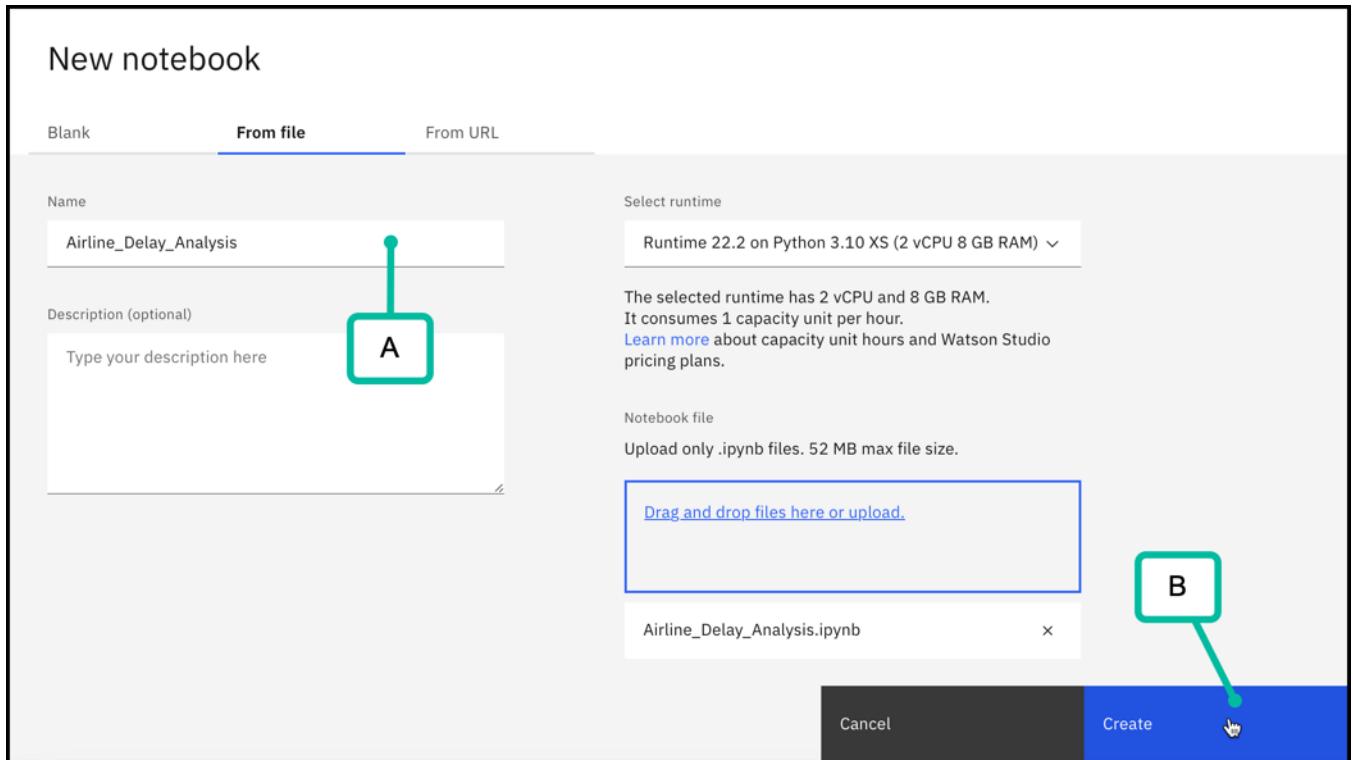
4. [A] Click on the **From file** tab of the *New notebook* screen. Then, [B] click on the **Drag and drop files here or upload** link to open a *Finder* window on Mac or a *File Explorer* window on Windows.



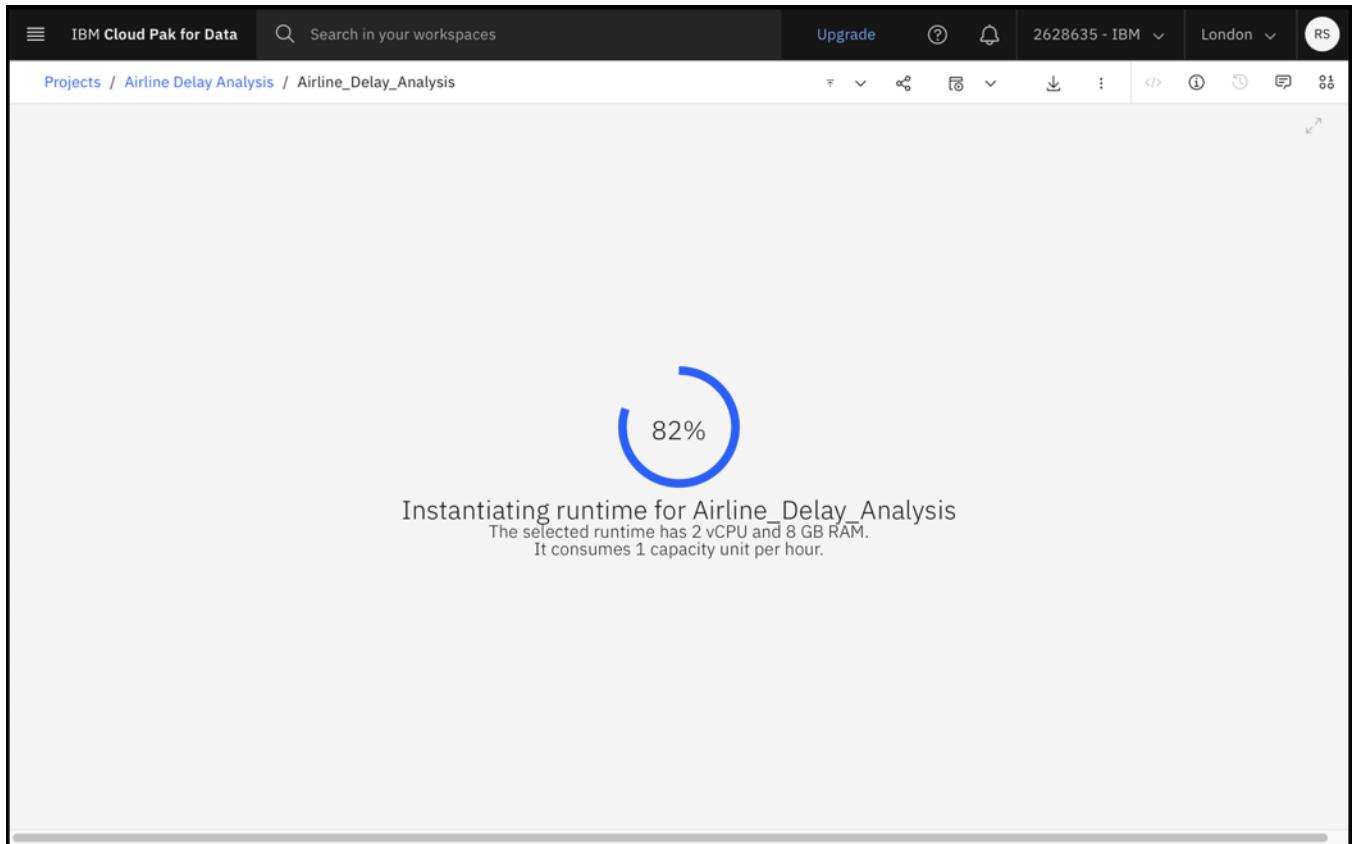
\_\_\_\_ 5. [A] Select the file named **Airline\_Delay\_Analysis.ipynb** and [B] click the **Upload** button.



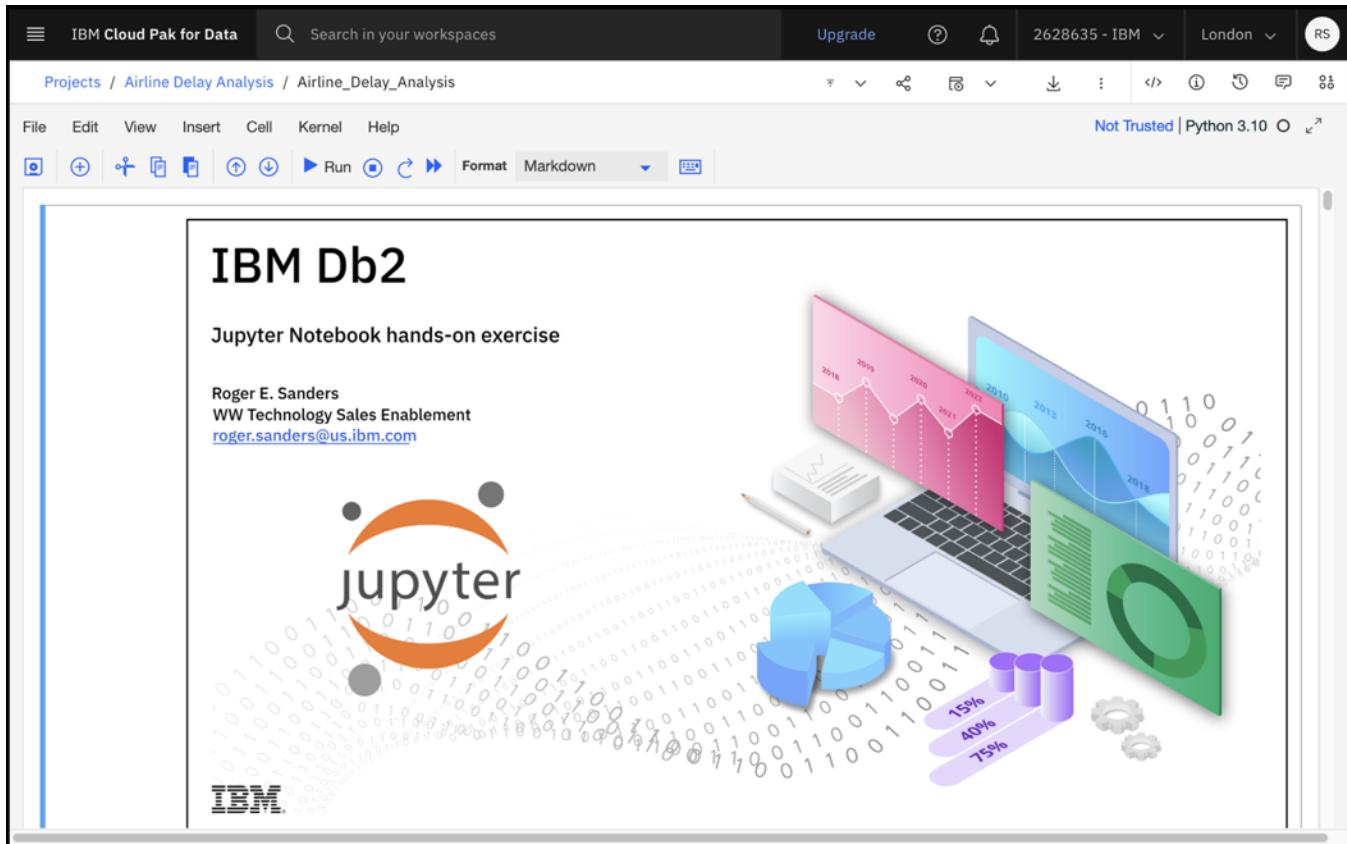
\_\_\_\_ 6. [A] Confirm that the **Name** field on the *New notebook* screen contains the name of the file just selected, minus the file extension (**Airline\_Delay\_Analysis**). Then, [B] click the **Create** button shown at the bottom of the screen.



7. The *New notebook* screen should be replaced with an *Instantiating runtime for [NotebookName]* screen that looks something like this:



\_\_\_\_ 8. Once the runtime for the Jupyter Notebook that was loaded has been instantiated, the *Instantiating runtime for [NotebookName]* screen should be replaced with a screen that looks similar to this:



The Jupyter Notebook is now ready for editing.

### 3. Prepare the Jupyter Notebook for execution.

\_\_\_\_ 1. Earlier, we learned that to work with an IBM Cloud resource from an application (such as a Jupyter Notebook), a service credential for that resource must exist. So, right after we created a Db2 on Cloud ‘Lite’ plan database, we created a service credential for that database (named **AirlineDelayDB**) and we copied the information associated with that service credential into a text file (named **AirlineDelayDB\_Credentials.txt**). Open the file that was saved earlier and locate the following information:

- [A] User name (“**username**”)
- [B] User password (“**password**”)
- [C] Database name (“**database**”)
- NOTE: The database name is always **bludb** (BLUDB) for Db2 on Cloud databases.*
- [D] Host name (“**hostname**”)
- [E] Port number (“**port**”)

```

    "environment": {},
    "type": "cli"
},
"db2": {
  "authentication": {
    "method": "direct",
    "password": "N6ogXjFSeE8XOJF@",
    "username": "vxp49176"
  },
  "certificate": {
    "certificate_base64":
      "LS0tLS1CRUdJTIBDRVJUSUZJQ0FURS0tLS0tCk1JSURIVENDQWdXZ0F3SUJBZ0lVT3dvMC9va09CUEN5RjFWefJxVGhKRW9ubDBV0RRWUpLb1pJaHZTkFRRUwKQIFBd0hqRWNNQm9HTFVRUF3d1RTVUpOSUVoc2lzVmJRVJoZEgDaViYTmxiekFIrnweU1EOTRNRF3TwpVMwpNaJphRncwek1EQTRNRREi3TwpVM01qWmFnQjR4SERBYUJnTiZCQU1NRTBsQ1RTQkrIzkkVnkNCRVIYUmhzbUZ6ClpYTxDunZ0VpTUewRONTcUdTSWlzRFFFQkFRVUFBNEICRHd8d2InRtBb0lCQVFEB0ZFNQ0SGdOeXZMUVlwR3gQKTBoamRXQnM4NVBjTDNyRSjN1R3K2d1RudQSUxJU0VZV3o4Y1g1TG1XQk0rY1FnOG9VeSrQXJ3OEoxaxDrZQpySmllU21clf4WTM0c3BQeGRFVEZkWEhScnJhMGU2VmM4MW42TlJl02HSn1Q3hrTG5GMUIFQW9hbHYwaDM2CnhDT0FvcXRwTfFrTzNmTRGeU0yRDRiajkxckl4RGk4V9XMVpVdVhMNGwzZXXVLZUVCeTruZmhJv3kySVc3aUMKbGpMZ3RIN3hZTDVHbVpKOUsdYWrtrSnj1cnpNREFQLzvUYnRIUlydEldotBTRNSVRFZHIESVFYUEZGRDBHYzloZao3M29JdpVZUJ3VC9uRHN80TJNC82SktzWpKN0lpdBFTN3Y2a2diUVhNDIBaUVJNxPQdUvpVzNOY9GR0p0YcmY2a2jBZ01CQUFHaiV6QJNQjBHQTFRVZERnUvdCQIR2zZ2RU5MRjFvbWzN0Q03MmxOcmMzSD12bURBUEJnTiZUk1CQWY4R0UJQURBUUgvTUEwRwpDU3FHU0l0M0RRRUJDd1VBQTRJQkFRQtgydFvnUTZlaTZWHZndDj0dJdrbkpva1Y5JWNkaTNzbFVFWkNDUytjCIVQZ3NmMnVBMDlxHiWTm1RkhjHZ1Vm0pVHRYTmk2NUN2WizSrnYxc3p1cu9zdFB5bkj4bIn4cUs0dk0dtkVkBWRUgcE11znZBSmxkV3c4UEJ7ZGJtTk1HdGM4SzlwT0050VdQb1ZFRXVXVGdEhJKTXFBZnpYXlidUV0dwpoCw1pV2swTmVXNGk5ZEY4S2dTWUVAQWFodXVBSIRldXB2R2RPV1U0eEV4bm03aEVrbmZPV2ZITThDd08xNWFZCIRGQ2s0Q0pDUmR4Mlg5U284V3o1Z3MzncryRKFDQlJyZONYyeFDZnZrZTUDvVNHNkxFRHJHbmpWaXVSQkpZdW4Kt1RxWXROaVBHaHpuTHJrL0Fzam1LmzBxQmFLTmFyNuDQajhqlapNb2RIZ04KLS0tLS1FTkQgQ0VSVEIGSUNBVEUts0tQo=",
    "name": "dd14d0c-1b52-4f63-a606-53ecba28771d"
  },
  "composed": [
    "db2://vxg49176:N6ogXjFSeE8XOJF@/2d46b6b4-cbf6-40eb-bbce-6251e6ba0300.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:32328/bludb?"
  ],
  "authSource": "admin&replicaSet=repset"
},
"database": "bludb",
"host_ros": [],
"hosts": [
  {
    "hostname": "2d46b6b4-cbf6-40eb-bbce-6251e6ba0300.bs2io90l08kqb1od8lcg.databases.appdomain.cloud",
    "port": 32328
  }
],
"jdbc_url": [
  "jdbc:db2://2d46b6b4-cbf6-40eb-bbce-6251e6ba0300.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:32328/bludb:user=<userid>;password=<your_password>;sslConnection=true;"
]

```

2. In the Jupyter Notebook that was just added to the project (**Airline\_Delay\_Analysis.ipynb**), **[A]** scroll down until you see the code section under the heading **Step 3a: Connect to a remote, Db2 on Cloud database**. Then, **[B]** click on this section to make it the current section. Next, copy the **hostname** value you found in the service credential file (**AirlineDelayDB\_Credentials.txt**) to the clipboard and paste it over the text “**replace-with-your-hostname**” found in this code section. (*Copy the value beside the **hostname** keyword in the service credential file and do NOT remove the quotation marks!*) Repeat this process for the **port** (“**replace-with-your-port-number**”), **username** (“**replace-with-your-userID**”), and **password** “**replace-with-your-password**” keywords. When you are finished, **[C]** click on the **Save Notebook** icon to make the changes made permanent. (NOTE: You can use the “hot keys” **command-C** on Mac or **Ctrl-C** on Windows to copy data to the clipboard; you can use the “hot keys” **command-V** on Mac or **Ctrl-V** on Windows to paste the copied data.)

**Step 3a: Connect to a remote, Db2 on Cloud database**

```
In [ ]: #-----#
# Initialize All User-Specific Connection Variables - Db2 on Cloud Database      #
#  IMPORTANT: UPDATE WITH VALUES FROM YOUR OWN ENVIRONMENT, AS PER LAB INSTRUCTIONS.  #
#-----#
# Define And Initialize The Appropriate Variables
dbName = "bludb"
hostName = "replace-with-your-hostname"
portNum = "replace-with-your-port-number"
userID = "replace-with-your-userID"
passWord = "replace-with-your-password"
secureComm = True           # Use SSL (Secure Sockets Layer) Communication

# Display A Status Message Indicating This Work Is Complete
print("\nUser-specific connection variable initialization work complete!\n")
```

**Step 3b: Connect to a local, on-premises Db2 database**

3. When finished, the code section under the heading **Step 3a: Connect to a remote, Db2 on Cloud database** should look something like this:

**Step 3a: Connect to a remote, Db2 on Cloud database**

```
In [ ]: #-----#
# Initialize All User-Specific Connection Variables - Db2 on Cloud Database      #
#  IMPORTANT: UPDATE WITH VALUES FROM YOUR OWN ENVIRONMENT, AS PER LAB INSTRUCTIONS.  #
#-----#
# Define And Initialize The Appropriate Variables
dbName = "bludb"
hostName = "2d46b6b4-cbf6-40eb-bbce-6251e6ba0300.bs2io90l08kqb1od8lcg.databases.appdomain.cloud"
portNum = "32328"
userID = "vxp49176"
passWord = "N6ogXjFSeE8X0JFo"
secureComm = True           # Use SSL (Secure Sockets Layer) Communication

# Display A Status Message Indicating This Work Is Complete
print("\nUser-specific connection variable initialization work complete!\n")
```

**Step 3b: Connect to a local, on-premises Db2 database**

4. Exit the Jupyter Notebook and return to the project by clicking on the **Airline Delay Analysis** project breadcrumb located at the top of the screen.

**Step 3a: Connect to a remote, Db2 on Cloud database**

```
In [ ]: #--#
# Initialize All User-Specific Connection Variables - Db2 on Cloud Database #
#   IMPORTANT: UPDATE WITH VALUES FROM YOUR OWN ENVIRONMENT, AS PER LAB INSTRUCTIONS. #
#--#
# Define And Initialize The Appropriate Variables
dbName = "bludb"
hostName = "2d46b6b4-cbf6-40eb-bbce-6251e6ba0300.bs2io90l08kqb1od8lcg.databases.appdomain.cloud"
portNum = "32328"
userID = "vxg49176"
passWord = "N6ogXjFSeE8X0JFo"
secureComm = True # Use SSL (Secure Sockets Layer) Communication
```

5. When the Assets screen for the **Airline Delay Analysis** project is displayed again, the **Airline\_Delay\_Analysis** Jupyter Notebook should be in the project assets list.

**Assets**

Name	Last modified	Actions
Airline_Delay_Analysis	1 hour ago Modified by you	⋮
Build_And_Populate_Table	1 day ago Modified by you	⋮
Airline_Delay_Cause_2013-20...	1 day ago Modified by you	⋮

**Congratulations!** You are now ready to begin working your way through the **Airline\_Delay\_Analysis** Jupyter Notebook.

## 4. Run the Airline\_Delay\_Analysis Jupyter Notebook

1. If you are not already there, navigate to the Assets screen for the **Airline Delay Analysis** project, and click on the link for the Jupyter Notebook asset named **Airline\_Delay\_Analysis**.

The screenshot shows the 'Assets' tab in the IBM Cloud Pak for Data interface. On the left, there's a sidebar with '3 assets' and categories for 'Data' (1 item) and 'Notebooks' (2 items). The main area displays a table titled 'All assets' with columns for 'Name', 'Last modified', and a 'More' (three-dot) menu. The assets listed are:

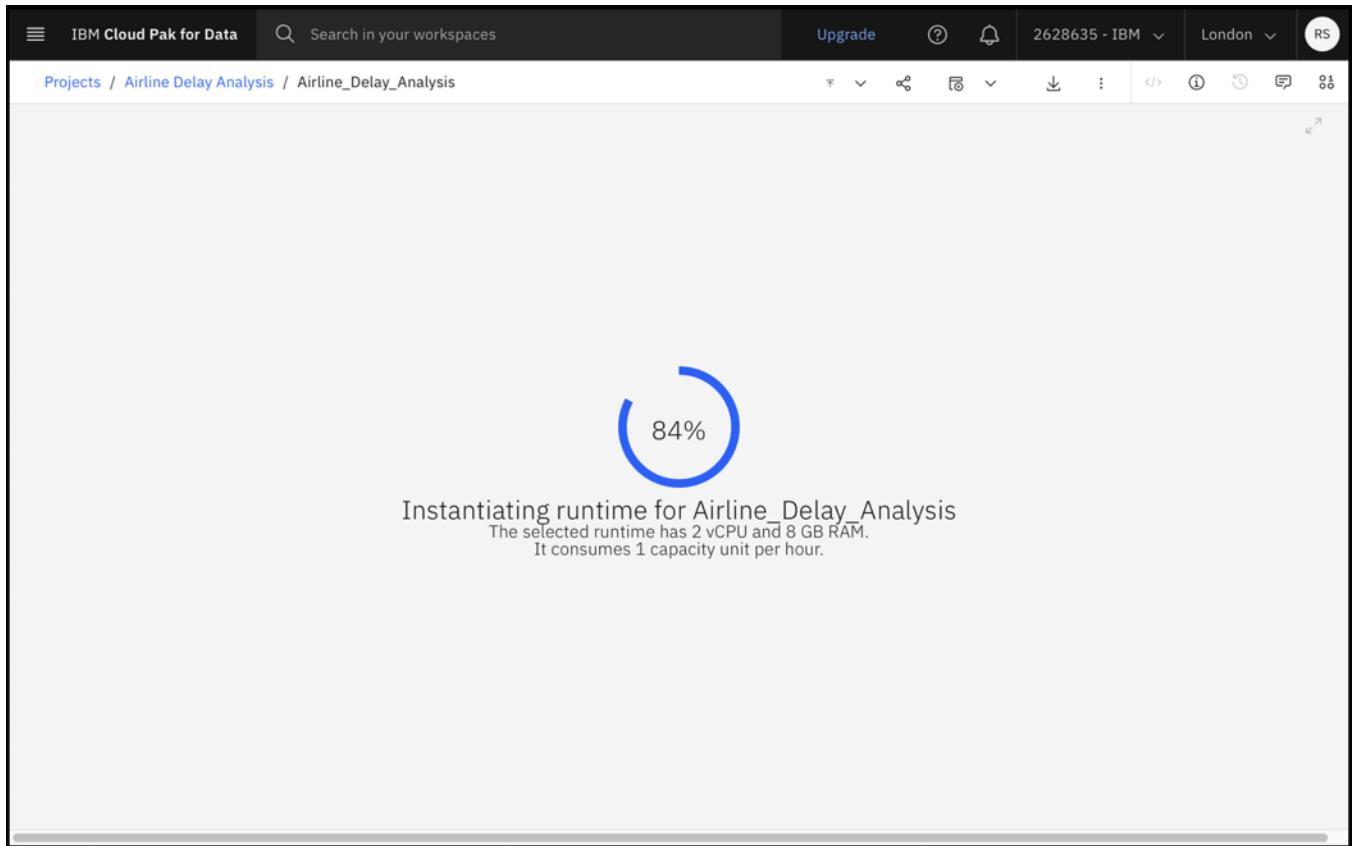
Name	Last modified	More
Airline_Delay_Analysis	4 hours ago Modified by you	⋮
Build_And_Populate_Table	1 week ago Modified by you	⋮
Airline_Delay_Cause_2013-20... CSV	1 week ago Modified by you	⋮

To the right, a panel titled 'Data in this project' contains a dashed box with the placeholder text 'Drop data files here or browse for files to upload'.

2. When the Jupyter Notebook opens, you may notice that the Jupyter Notebook menu and icons are not available. That's because the notebook has not been instantiated (i.e., a kernel for the Jupyter Notebook has not been started). Instantiate the **Airline\_Delay\_Analysis** Jupyter Notebook by clicking on the **Edit** (pencil) icon located at the top of the screen.

The screenshot shows a Jupyter Notebook interface with the title 'IBM Db2'. The top bar includes standard Jupyter Notebook icons like 'File', 'Edit', 'Cell', 'Kernel', 'Help', and 'Run'. The 'Edit' icon is highlighted with a mouse cursor. The notebook content itself is titled 'Jupyter Notebook hands-on exercise' and includes author information: 'Roger E. Sanders', 'WW Technology Sales Enablement', and an email address 'roger.sanders@us.ibm.com'. There's also a small 'jupyter' logo at the bottom left. The right side of the screen features a decorative graphic with a laptop, a pencil, and various charts and binary code.

3. While the Jupyter Notebook is being instantiated, you should see a screen that looks something like this:



\_\_\_\_ 4. Once the **Airline\_Delay\_Analysis** Jupyter Notebook has been instantiated, work your way through it, one cell at a time, reading the documentation found in markup cells and executing the source code found in code cells. The notebook is well documented and easy to work through. As a reminder, to execute a code cell, simply click on the cell to select it and then click on the ► Run icon.

The screenshot shows a Jupyter Notebook interface within the IBM Cloud Pak for Data environment. The top navigation bar includes 'IBM Cloud Pak for Data', a search bar, 'Upgrade', user information, and location settings ('London'). The main area displays a project titled 'Airline Delay Analysis / Airline\_Delay\_Analysis'. The toolbar below the menu bar includes icons for file operations, cell selection, and execution. A tooltip 'Run Cell, Select Below' points to the 'Run' button. The content area contains a section header 'Section 1. Prepare the lab environment' and a sub-section 'Step 1. Download and install the appropriate software packages'. It includes an 'Overview' note about installing Python packages and a 'Execute the code:' section with instructions and a code cell. The code cell content is as follows:

```
In [ ]: #-----#
# Install The Appropriate Software Packages #
#
# NOTE: The code in this section only needs to be executed once in a runtime environment and #
# the packages may have already been installed. If so, it is not harmful to attempt to install #
# the packages identified again, as subsequent attempts will simply state that the package #
# requirement has already been satisfied. #
#
#-----#
#-----#
# Download And Install The ibm db Driver Packae #
```

\_\_\_\_ 5. You will find that with this Jupyter Notebook, some of the code cells generate a significant amount of output. When that is the case, use the scroll bars that are provided with the output window to view all the information produced.

IBM Cloud Pak for Data

Search in your workspaces

Upgrade ③ ⚡ 2628635 - IBM London RS

Projects / Airline Delay Analysis / Airline\_Delay\_Analysis

File Edit View Insert Cell Kernel Help Not Trusted | Python 3.10

```
# Display A Status Message Indicating This Work Is Complete
#
#-----#
print("All software packages needed have been installed!\n")

Requirement already satisfied: pytz>=2020.1 in /opt/conda/envs/Python-3.10/lib/python3.10/site-packages (from pandas>=0.23->seaborn) (2022.1)
Requirement already satisfied: six>=1.5 in /opt/conda/envs/Python-3.10/lib/python3.10/site-packages (from python-dateutil>=2.7->matplotlib>=2.2->seaborn) (1.16.0)

--2023-06-07 20:52:38-- https://raw.githubusercontent.com/IBM/db2-jupyter/master/db2.ipynb
Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.109.133, 185.199.108.133, 185.199.111.133, ...
...
Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|185.199.109.133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 130924 (128K) [text/plain]
Saving to: 'db2.ipynb'

db2.ipynb      100%[=====] 127.86K --.-KB/s   in 0.004s

2023-06-07 20:52:39 (29.5 MB/s) - 'db2.ipynb' saved [130924/130924]

All software packages needed have been installed!
```

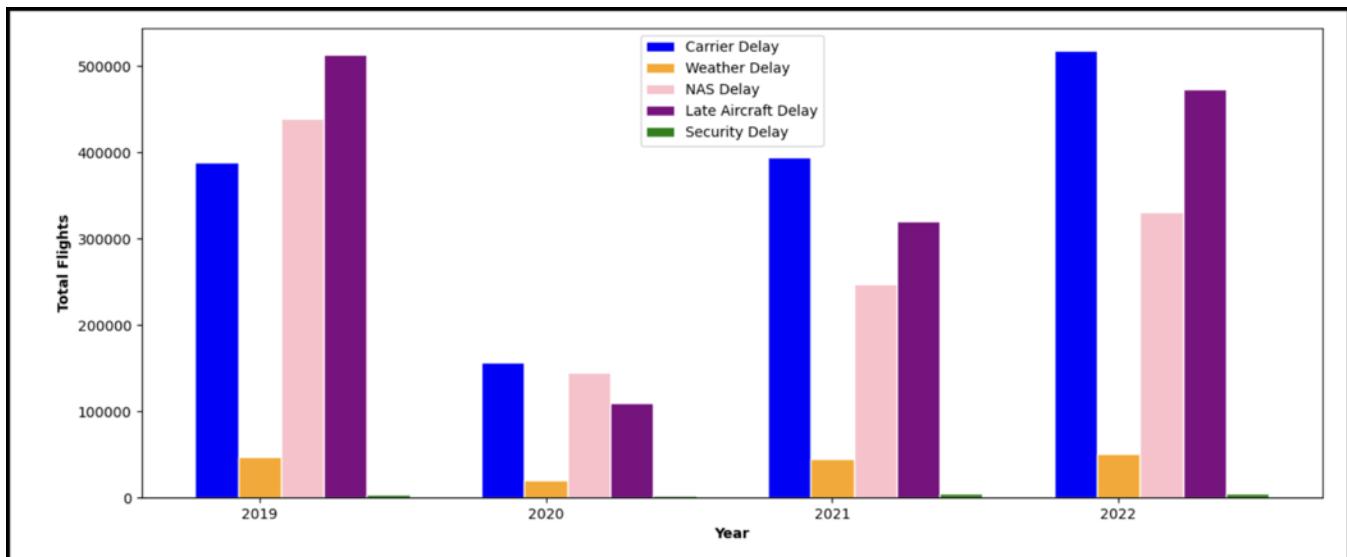
## Step 2. Set up the Jupyter Notebook environment

**Overview:**

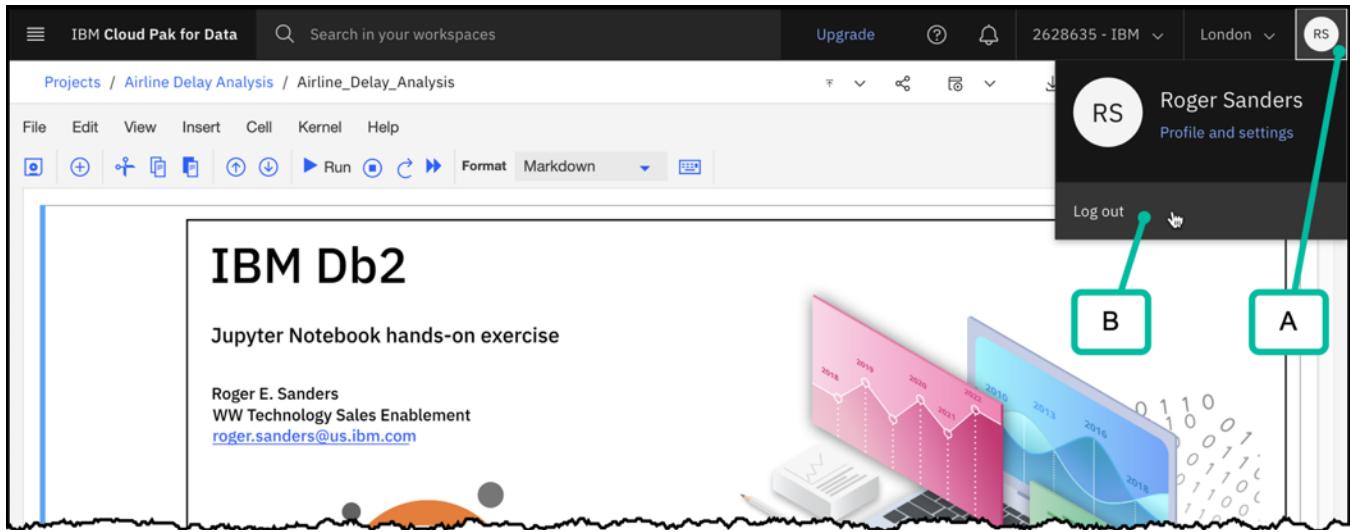
Before you can begin interacting with an IBM Db2 database using Python or Jupyter Notebook, there are some basic steps you must perform. These steps include:

1. Loading (importing) the `ibm_db` driver into the Python application or Jupyter Notebook.

6. Pay close attention to the output produced by the code cells that generate graphs. These graphs are designed to help you visualize the data analysis that is performed.



7. When you are finished, log out of IBM Cloud Pak for Data as a Service by **[A]** clicking on your user icon located in the top right-hand corner of the screen and then **[B]** clicking on the **Log out** item in the menu presented.



**Congratulations!** You have just used Watson Studio (running in IBM Cloud Pak for Data as a Service) and a Jupyter Notebook to analyze data stored in a Db2 on Cloud database.

## VI. Thank you!

Thank you for participating in this workshop! I hope you found it both enjoyable and informative.