



Cloud Pak for Data
Version 2.5 or higher
Tutorial – Mortgage

Contents

1.	Prerequisites	5
2.	Setting up database and sample data.....	5
3.	Access Credentials	6
3.1.	Access credential for Db2 database.....	6
3.2.	Sign in to Cloud Pak for Data web console as Administrator.....	6
4.	Create Connection	7
4.2.	Navigate to Connections	7
4.3.	Add connection	7
5.	Discover Assets	8
5.1.	Navigate to discover assets	8
6.	Add users	11
7.	Implement Business Glossary.....	14
7.1.	Download Business Glossaries.....	14
7.2.	Import Categories	15
7.2.	Import Terms	16
7.3.	Create a policy	17
7.4.	Create a rule	17
7.5.	Add rule to metadata.....	19
8.	Access data as a Data Scientist	21
8.1.	Create analytic project.....	21
8.2.	Assets from Glossary	22
8.3.	Check Asset Details.....	22
8.4.	Global search	Error! Bookmark not defined.
10.	Navigate to data catalog.....	Error! Bookmark not defined.
11.	Data Virtualization.....	24
11.1.	Adding a new data source for Db2	24
11.5.	Add virtual table to catalog	27
11.6.	Publish virtualized table.....	27
11.7.	Access information for virtual table	28
11.8.	Deliver Dataset	Error! Bookmark not defined.
12.	Build Model	29
12.1.	Navigate to analytics project	29

Cloud Pak for Data (v.2.5 or higher) – Tutorial

12.2. Create deployment space	29
12.3. Create notebook	30
12.4. Review and run notebook.....	30
12.5. Test the model.....	32

Cloud Pak for Data is a single end to end platform for data management, governance and data science analytics. It provides a one stop shop for data scientists, data engineer and data stewards to collaborate on the platform to acquire, govern and extract best insights from the data in the least amount of time.

In this demo, user will use a set of a fictitious mortgage data that available in Db2 database on a docker image. User will perform following tasks to predict if a prospective customer may default on their mortgage.

- Create connection from Cloud Pak for Data to Db2 database on cloud
- Discover Db2 assets from Cloud Pak for Data
- Transform the Db2 data using Data Virtualization
- Build a simple machine learning model for prediction

1. Prerequisites

- Access to an operational Cloud Pak for Data (v.2.5 or higher) Instance
- Install Git on the machine that you will use for the tutorial

2. Setting up database and sample data

Log in to the cluster where Cloud Pak for Data is deployed or log in to a Linux-based system (RedHat or Ubuntu) that can access the cluster over your network.

From your home directory, clone the tutorial sample files:

```
git clone https://github.com/IBM-ICP4D/icp4d-tutorials.git
```

Change to the tutorials directory:

```
cd icp4d-tutorials/tutorials/
```

The sample data-loading utility, load_samples.sh, provides an easy way to host a Db2 server and load it with sample data.

Run the following command to view the list of sample data that is provided in the load_samples.sh utility:

```
./load_samples.sh -l
```

Run the following command to load the sample data into a Db2 database:

```
./load_samples.sh -t mortgage-002
```

After the loading process completes, an instance of Db2 is hosted on your cluster as a Docker container.

3. Access Credentials

To work through the tutorial, you need access a Db2 database.

3.1. Access credential for Db2 database

For this tutorial you need JDBC connection to access to a Db2 database that hosted locally on Cloud Pak for Data. Following are JDBC connection credential for Db2:

JDBC Host name	<Same IP address as your web console>
Port number	50000
Database name	MORTGAGE
User ID	db2inst1
Password	password
Db2	Version 11.1
JDBC connection string	jdbc:db2://<same IP as Web Console>:50000/MORTGAGE

3.2. Sign into Cloud Pak for Data web console as Administrator

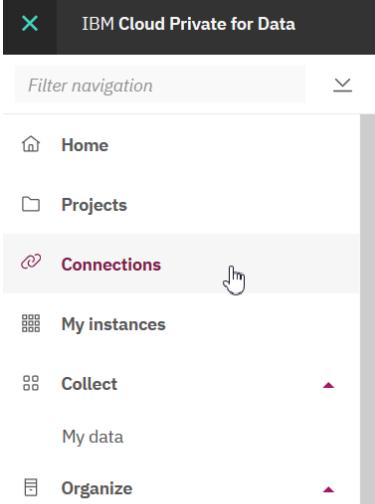
You should have an operational Cloud Pak for Data Instance. Use latest version of Firefox or Google Chrome browser to access the Cloud Pak for Data web console. Starting from here all instruction need to execute on Cloud Pak for Data web console only. You need to login as admin who has administrator privileges.

Sign in  Username <input type="text" value="admin"/> PASSWORD <input type="password" value="*****"/> <input type="button" value="Sign In"/>	<p>Sigh into the Cloud Pak for Data web console as user ‘admin’ and password is ‘password’.</p>
--	---

4. Create Connection

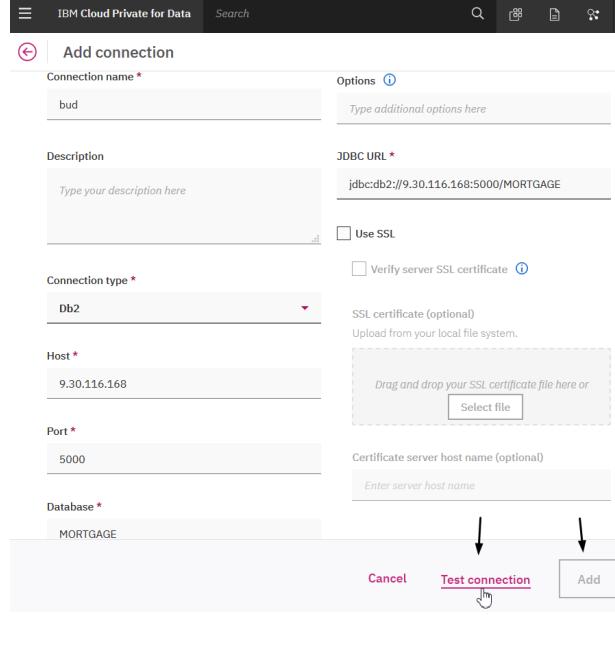
Create a connection to the data source for Db2 database.

4.2. Navigate to Connections



On the left pane choose **Connections**. Next, on the **Data Connections** window click on the  icon.

4.3. Add connection



Fill out the **Add Connection** information according to the information provided in step 3.1. Access credential for DB2. Credential used in following step is just an example.

1. For **Choose connection** use the drop-down menu and select ‘Db2’.
2. Use ‘Bud’ as the **Connection name**
3. Use IP of the cluster node (where DB2 database) as **Host**
4. **Port** is ‘50000’
5. **Database is** ‘MORTGAGE’
6. **Username** is ‘db2inst1’ and **Password** is ‘password’.

Next click on **Test Connection**, once it successful click on **Save Connection**.

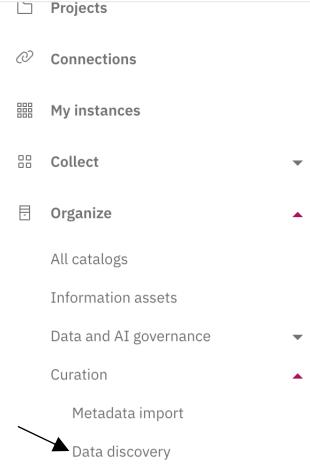


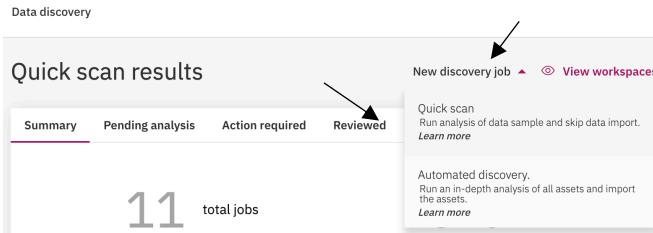
Success The test connection was successful. Click **Add** to save the connection information.

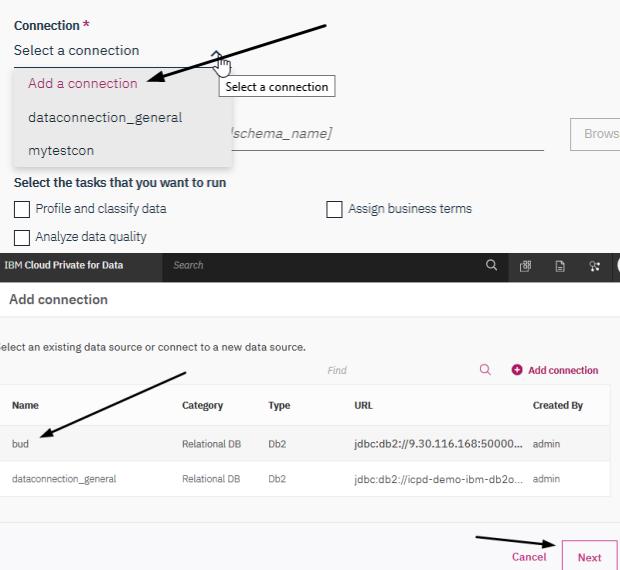
5. Discover Assets

Use the data source created above discover all data assets from Db2 database.

5.1. Navigate to discover assets

	<p>From Organize option on the left pane, choose Curation > Data discovery.</p>
---	--

	<p>To select discover job</p> <p>Navigate to New discover job > Quick scan</p>
--	--

	<p>To discover assets</p> <ol style="list-style-type: none"> 1. Click on Add a connection 2. Choose the connection named bud that you created previously, click Next
---	---

Quick scan job

Connection * bud

Discovery root ⓘ schema[MORTGAGE|DB2INST1]

Discovery options

- Analyze columns
- Analyze data quality
- Assign terms
- Use machine learning to assign terms
- Use data sampling

The maximum number of records included in the data set sample: 1000

Workspace * ⓘ Mortgage

3. Choose the connection named **bud** that you created previously.
4. Select **Discover root** as **MORTGAGE > DB2INST1**
5. Check necessary **Discover options**
6. Click on **Add a workspace** under Workspace and named it as **Mortgage**. Click **Create**.
7. Click on **Discover**

It may take few minutes to complete.

Click on **Quick scan results** > **Action required** > **View results** or **View workspaces** to explore the discover assets.

Quick scan results

New discovery job ⓘ View workspaces ⓘ View automated discovery results

Summary	Pending analysis	Action required	Reviewed													
Status <ul style="list-style-type: none"> <input checked="" type="radio"/> All jobs pending analysis <input type="radio"/> Analyzing <input type="radio"/> In queue for analysis 	Pause View results	1 item selected (select up to 15) Cancel														
	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Job ID</th> <th style="text-align: left;">Data assets</th> <th style="text-align: left;">Connection</th> <th style="text-align: left;">Started by</th> <th style="text-align: left;">Processing time</th> <th style="text-align: left;">Status</th> <th style="text-align: left;">Status updated</th> </tr> </thead> <tbody> <tr> <td>qs_1571071613091</td> <td>-</td> <td>bud</td> <td>admin</td> <td>2 minutes 15 seconds</td> <td>Analyzing</td> <td>-</td> </tr> </tbody> </table>	Job ID	Data assets	Connection	Started by	Processing time	Status	Status updated	qs_1571071613091	-	bud	admin	2 minutes 15 seconds	Analyzing	-	
Job ID	Data assets	Connection	Started by	Processing time	Status	Status updated										
qs_1571071613091	-	bud	admin	2 minutes 15 seconds	Analyzing	-										

Results f qs_1571071613091

Explore assets

Business term assignment		Data class assignment	
Assigned	100	Assigned	70.37
Suggested		Suggested	51.85
Unassigned		Unassigned	29.63

Review the discovery results using **Explore assets** tab

<p>Asset type</p> <ul style="list-style-type: none"> <input type="radio"/> File <input type="radio"/> Schema <input checked="" type="radio"/> Table <input checked="" type="radio"/> Column <p>Filters</p> <p>Labels</p> <p>No filters of this type</p> <p>Tables 3 X</p> <p><input checked="" type="checkbox"/> 3 tables selected</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> MORTGAGE_CUSTOMER <input checked="" type="checkbox"/> MORTGAGE_PROPERTY <input checked="" type="checkbox"/> MORTGAGE_DEFAULT <input type="checkbox"/> MORTGAGE_JOIN <p>Clear Apply</p>	<p>Discovered columns (27)</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Column name</th> <th>Identity</th> <th>Quality</th> <th>Assigned business term</th> <th>Suggested business term</th> <th>Assigned data class</th> <th>Suggested data class</th> <th>Business term actions</th> </tr> </thead> <tbody> <tr><td>APPLIED_ONLINE</td><td>MORTGAGE_JOIN</td><td>100%</td><td>-</td><td>-</td><td>Indicator 100%</td><td>-</td><td></td></tr> <tr><td>APPLIED_ONLINE</td><td>MORTGAGE_CUSTOMER</td><td>100%</td><td>-</td><td>-</td><td>Indicator 100%</td><td>-</td><td></td></tr> <tr><td>CARD_DEBT</td><td>MORTGAGE_CUSTOMER</td><td>96%</td><td>-</td><td>-</td><td>-</td><td>US Zip Code 2%</td><td></td></tr> <tr><td>CARD_DEBT</td><td>MORTGAGE_JOIN</td><td>96%</td><td>-</td><td>-</td><td>-</td><td>US Zip Code 4%</td><td></td></tr> <tr><td>CURRENT_LOANS</td><td>MORTGAGE_JOIN</td><td>100%</td><td>-</td><td>-</td><td>Boolean 100%</td><td>Indicator 100%</td><td></td></tr> <tr><td>CURRENT_LOANS</td><td>MORTGAGE_CUSTOMER</td><td>100%</td><td>-</td><td>-</td><td>Boolean 100%</td><td>Indicator 100%</td><td></td></tr> <tr><td>ID</td><td>MORTGAGE_DEFAULT</td><td>100%</td><td>-</td><td>-</td><td>Identifier 100%</td><td>-</td><td></td></tr> <tr><td>ID</td><td>MORTGAGE_CUSTOMER</td><td>100%</td><td>-</td><td>-</td><td>Identifier 100%</td><td>-</td><td></td></tr> <tr><td>ID</td><td>MORTGAGE_PROPERTY</td><td>100%</td><td>-</td><td>-</td><td>Identifier 100%</td><td>-</td><td></td></tr> </tbody> </table> <p>Items per page: 10 1–10 of 27 items</p> <p>1 of 3 pages </p>	Column name	Identity	Quality	Assigned business term	Suggested business term	Assigned data class	Suggested data class	Business term actions	APPLIED_ONLINE	MORTGAGE_JOIN	100%	-	-	Indicator 100%	-		APPLIED_ONLINE	MORTGAGE_CUSTOMER	100%	-	-	Indicator 100%	-		CARD_DEBT	MORTGAGE_CUSTOMER	96%	-	-	-	US Zip Code 2%		CARD_DEBT	MORTGAGE_JOIN	96%	-	-	-	US Zip Code 4%		CURRENT_LOANS	MORTGAGE_JOIN	100%	-	-	Boolean 100%	Indicator 100%		CURRENT_LOANS	MORTGAGE_CUSTOMER	100%	-	-	Boolean 100%	Indicator 100%		ID	MORTGAGE_DEFAULT	100%	-	-	Identifier 100%	-		ID	MORTGAGE_CUSTOMER	100%	-	-	Identifier 100%	-		ID	MORTGAGE_PROPERTY	100%	-	-	Identifier 100%	-		<p>Review assets for proper business data class assignment, if needed you can adjust them.</p> <p>Select Asset type as “Column”</p> <p>Filters necessary tables using checkbox</p> <p>Click on Apply</p>
Column name	Identity	Quality	Assigned business term	Suggested business term	Assigned data class	Suggested data class	Business term actions																																																																											
APPLIED_ONLINE	MORTGAGE_JOIN	100%	-	-	Indicator 100%	-																																																																												
APPLIED_ONLINE	MORTGAGE_CUSTOMER	100%	-	-	Indicator 100%	-																																																																												
CARD_DEBT	MORTGAGE_CUSTOMER	96%	-	-	-	US Zip Code 2%																																																																												
CARD_DEBT	MORTGAGE_JOIN	96%	-	-	-	US Zip Code 4%																																																																												
CURRENT_LOANS	MORTGAGE_JOIN	100%	-	-	Boolean 100%	Indicator 100%																																																																												
CURRENT_LOANS	MORTGAGE_CUSTOMER	100%	-	-	Boolean 100%	Indicator 100%																																																																												
ID	MORTGAGE_DEFAULT	100%	-	-	Identifier 100%	-																																																																												
ID	MORTGAGE_CUSTOMER	100%	-	-	Identifier 100%	-																																																																												
ID	MORTGAGE_PROPERTY	100%	-	-	Identifier 100%	-																																																																												

<p>Asset type</p> <ul style="list-style-type: none"> <input type="radio"/> File <input type="radio"/> Schema <input checked="" type="radio"/> Table <input type="radio"/> Column <p>Filters</p> <p>Schemas</p> <p><input type="checkbox"/> DB2INST1</p> <p>Status</p>	<p>Approve results Reject results Audit assets</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Table name</th> <th>Identity</th> <th>Quality</th> <th>Schema name</th> <th>Discovery root</th> <th>Status</th> </tr> </thead> <tbody> <tr><td><input checked="" type="checkbox"/> MORTGAGE_CUSTOMER</td><td>MORTGAGE_CUSTOMER_qs_1573834179481</td><td>100%</td><td>DB2INST1</td><td>schema[MORTGAGE DB:</td><td>Ready for review</td></tr> <tr><td><input checked="" type="checkbox"/> MORTGAGE_DEFAULT</td><td>MORTGAGE_DEFAULT_qs_1573834179481</td><td>100%</td><td>DB2INST1</td><td>schema[MORTGAGE DB:</td><td>Ready for review</td></tr> <tr><td><input type="checkbox"/> MORTGAGE_JOIN</td><td>MORTGAGE_JOIN_qs_1573834179481</td><td>100%</td><td>DB2INST1</td><td>schema[MORTGAGE DB:</td><td>Ready for review</td></tr> <tr><td><input checked="" type="checkbox"/> MORTGAGE_PROPERTY</td><td>MORTGAGE_PROPERTY_qs_1573834179481</td><td>100%</td><td>DB2INST1</td><td>schema[MORTGAGE DB:</td><td>Ready for review</td></tr> </tbody> </table>	Table name	Identity	Quality	Schema name	Discovery root	Status	<input checked="" type="checkbox"/> MORTGAGE_CUSTOMER	MORTGAGE_CUSTOMER_qs_1573834179481	100%	DB2INST1	schema[MORTGAGE DB:	Ready for review	<input checked="" type="checkbox"/> MORTGAGE_DEFAULT	MORTGAGE_DEFAULT_qs_1573834179481	100%	DB2INST1	schema[MORTGAGE DB:	Ready for review	<input type="checkbox"/> MORTGAGE_JOIN	MORTGAGE_JOIN_qs_1573834179481	100%	DB2INST1	schema[MORTGAGE DB:	Ready for review	<input checked="" type="checkbox"/> MORTGAGE_PROPERTY	MORTGAGE_PROPERTY_qs_1573834179481	100%	DB2INST1	schema[MORTGAGE DB:	Ready for review	<p>Change Asset type as “Table”</p> <p>Select all Mortgage related tables</p> <p>Click on Approve results</p>
Table name	Identity	Quality	Schema name	Discovery root	Status																											
<input checked="" type="checkbox"/> MORTGAGE_CUSTOMER	MORTGAGE_CUSTOMER_qs_1573834179481	100%	DB2INST1	schema[MORTGAGE DB:	Ready for review																											
<input checked="" type="checkbox"/> MORTGAGE_DEFAULT	MORTGAGE_DEFAULT_qs_1573834179481	100%	DB2INST1	schema[MORTGAGE DB:	Ready for review																											
<input type="checkbox"/> MORTGAGE_JOIN	MORTGAGE_JOIN_qs_1573834179481	100%	DB2INST1	schema[MORTGAGE DB:	Ready for review																											
<input checked="" type="checkbox"/> MORTGAGE_PROPERTY	MORTGAGE_PROPERTY_qs_1573834179481	100%	DB2INST1	schema[MORTGAGE DB:	Ready for review																											

<p>Approve assets</p> <p>The selected assets will be added to the catalog so that other users can access them.</p> <p>The analysis results for these assets will not be included in the catalog until you publish them.</p> <p>The analysis results will be loaded to the workspace that you selected when you started the new discovery job. In the workspace, you can run further analysis, edit the results, or publish them.</p> <p>Selected assets</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <input checked="" type="checkbox"/> MORTGAGE_CUSTOMER <input checked="" type="checkbox"/> MORTGAGE_DEFAULT <input checked="" type="checkbox"/> MORTGAGE_PROPERTY </div>	<p>Click on Approve assets</p>
---	--

6. Add users

Create users with different roles.

<p>Home</p> <p>Projects</p> <p>Connections</p> <p>My instances</p> <p>Collect</p> <p>Organize</p> <p>Analyze</p> <p>Administer</p> <ul style="list-style-type: none"> Manage platform Configure platform Gather diagnostics Manage users 	<p>From Administer option on the left pane, choose Manage users.</p>
---	--

<p>IBM Cloud Private for Data</p> <p>Manage users</p> <p>Users Roles</p> <p>Add user Connect to an LDAP server</p>	<p>Switch tab to ‘Users’ and click on ‘Add user’</p>
--	--

<p>New user</p> <p>User * dst1</p> <p>Username * dst1</p> <p>Password * *****</p> <p>Re-enter new password * *****</p> <p>Email * dst1@mail.com</p> <p>Roles *</p> <ul style="list-style-type: none"> <input type="checkbox"/> Administrator <input type="checkbox"/> Business Analyst <input type="checkbox"/> Data Engineer <input type="checkbox"/> Data Quality Analyst <input checked="" type="checkbox"/> Data Scientist <input type="checkbox"/> Data Steward <input type="checkbox"/> Developer <p>Cancel Save</p>	<p>Fill out Add User information for a data scientist</p> <ol style="list-style-type: none"> 1. User as dst1 2. Username is dst1 3. Use a valid email address 4. Set Password as dst1 5. Chose the user roles as Data Scientist <p>Click on Save to confirm the add user</p>
---	---

Follow same steps in Add User section (above) and two more account. Create **deng1** for Data Engineer and **dstw1** a data steward.

User	Role	Password
• deng1	Data Engineer	deng1
• dstw1	Data Stewards	dstw1

Log out from user **admin**

6.1. Grant Catalog Permission

With Watson Knowledge Catalog, you use catalogs to easily find and share your data and other assets. A catalog is like a private community for your organization. It's a way to organize resources for many data science projects: data assets, analytical assets, and the users who need to use the assets. You can manage access to the catalog by adding collaborators with specific roles that determine their permissions to perform actions.

Go to **Organize > All catalogs**

Select **Default Catalog** and click on action icon

Choose **View**

Go to **Access Control** tab

Click on **Add Collaborator**

	<p>Add Collaborators as Viewers</p> <p>Collaborators (0)</p> <p>d</p> <table border="1"><tr><td>admin</td><td>--</td></tr><tr><td>deng1</td><td>deng1@mail.com</td></tr><tr><td>dst1</td><td>dst1@mail.com</td></tr><tr><td>dstw1</td><td>dstw1@mail.com</td></tr></table> <p>Having trouble adding collaborators?</p> <p>Cancel Add</p>	admin	--	deng1	deng1@mail.com	dst1	dst1@mail.com	dstw1	dstw1@mail.com	<p>Add ‘deng1’ user as collaborator with editor role</p> <p>Add ‘dstw1’ user as collaborator with editor role</p> <p>Add ‘dst1’ user as collaborator with viewer role</p>
admin	--									
deng1	deng1@mail.com									
dst1	dst1@mail.com									
dstw1	dstw1@mail.com									

7. Implement Business Glossary

Cloud Pak for Data enables you to structure your enterprise information in a logical way, discover relationships between assets, and keep your data always up-to-date. You can import existing glossary with categories, terms, information governance policies and rules.

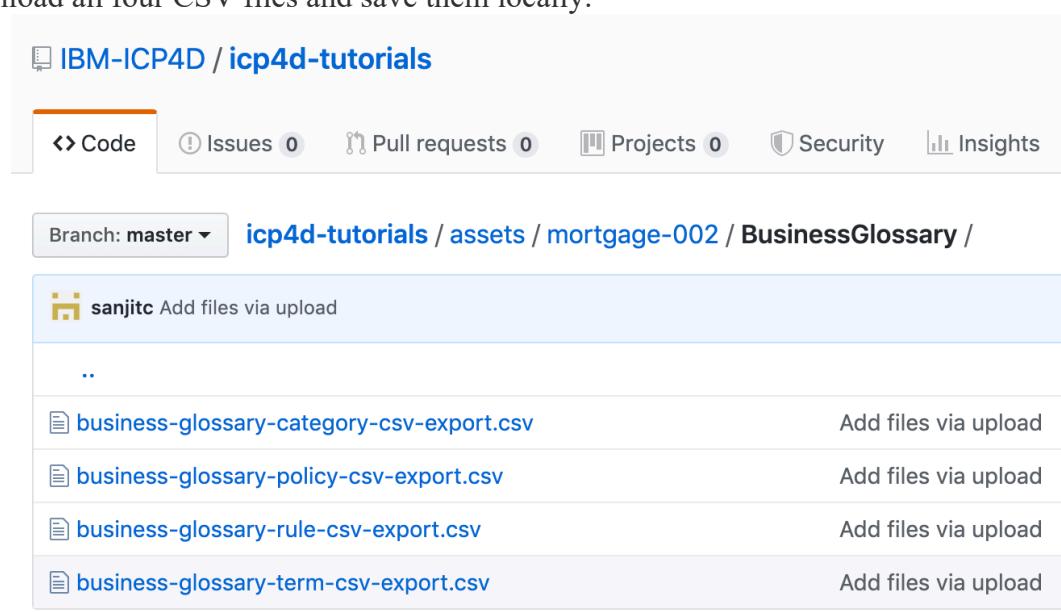
	<p>Sigh into the Cloud Pak for Data web console as user ‘dstw1’ and password is ‘dstw1’ that you created earlier.</p>
---	---

7.1. Download Business Glossaries

First download business glossaries from the GIT to your local machine.

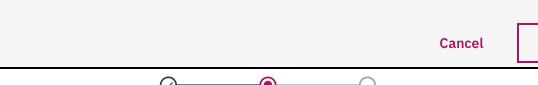
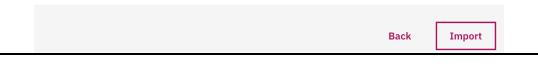
Go to: <https://github.com/IBM-ICP4D/icp4d-tutorials/tree/master/assets/mortgage-002/BusinessGlossary>

Download all four CSV files and save them locally.



7.2. Import Categories

Sequence is important when importing business glossaries. Make sure import categories before do the terms.

<p>Choose Organize > Data and AI governance > Categories from the left pane.</p> 	
	<p>Click on Import to import the CSV file contains category information that you downloaded from Git.</p>
 <p>Choose file</p> <p>Must be a CSV file.</p> <p>business-glossary-category-csv-export.csv</p> <p>The CSV file must conform to the template for importing governance artifacts. Learn more</p>	<p>Choose the CSV file location</p> <p>Click Next</p>
 <p>Select merge option</p> <ul style="list-style-type: none"> <input checked="" type="radio"/> Replace all values Imported values in the CSV file replace existing values in the catalog. <input type="radio"/> Replace with defined values Imported CSV values that are not empty replace existing values in the catalog. <input type="radio"/> Replace empty values Imported values in the CSV file replace only empty values in the catalog. 	<p>Select merge option as Replace all values</p> <p>Click Import</p>
	

7.2. Import Terms

Choose **Organize > Data and AI governance > Business terms** from the left pane.



Click on **Import** to import the CSV file contains term information that you downloaded from Git.

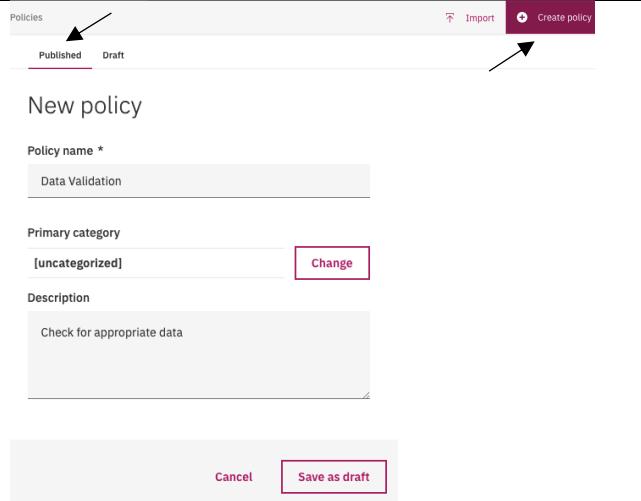
<p>The screenshot shows the first step of the import wizard: 'Choose file'. It has three circular steps: 'Choose file' (with a red circle), 'Set merging' (empty), and 'Import' (empty). Below the steps, the text 'Choose file' is bolded. A file selection dialog is open, showing a single file: 'business-glossary-term-csv-export.csv'.</p>	<p>Choose the CSV file location</p> <p>Click Next</p>
<p>The screenshot shows the second step of the import wizard: 'Select merge option'. It has three circular steps: 'Choose file' (with a red circle), 'Set merging' (with a red circle), and 'Import' (empty). Below the steps, the text 'Select merge option' is bolded. There are three radio button options: <ul style="list-style-type: none"> Replace all values: Imported values in the CSV file replace existing values in the catalog. Replace with defined values: Imported CSV values that are not empty replace existing values in the catalog. Replace empty values: Imported values in the CSV file replace only empty values in the catalog. A 'Next' button is highlighted with a red box at the top right of the dialog. </p>	<p>Select merge option as Replace all values</p> <p>Click Import</p> <p>Review each imported business terms and then publish</p>

7.3. Create a policy

Create governance policies and rules for the entire organization to ensure clarity and compatibility among departments, projects, or products.

Choose **Organize > Data and AI governance > Policy** from the left pane

Select **Published** tab and click on **Create Policy**

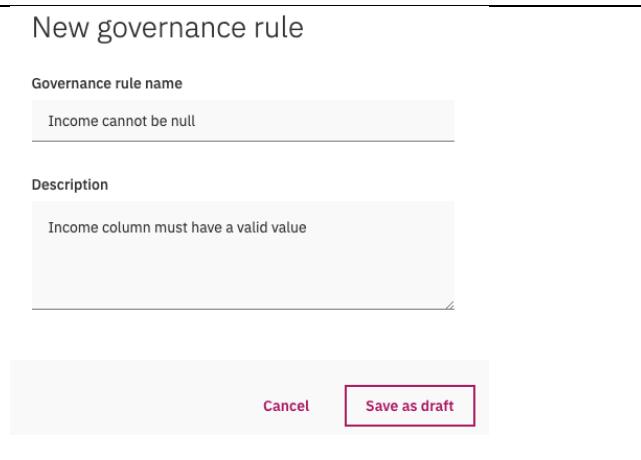
 <p>New policy</p> <p>Policy name * Data Validation</p> <p>Primary category [uncategorized] Change</p> <p>Description Check for appropriate data</p> <p>Cancel Save as draft</p>	<p>On the New policy window create a policy with following information and click on Save as draft:</p> <p>Name: Data Validation Description: Check for appropriate data</p> <p>It will take few minutes to appear under list of available policies.</p> <p>Once new policy available let's publish it.</p>
--	--

7.4. Create a rule

Choose **Organize > Data and AI governance > Rule** from the left pane

Select **Published** tab and click on **Create Rule**

Choose **Governance rule**

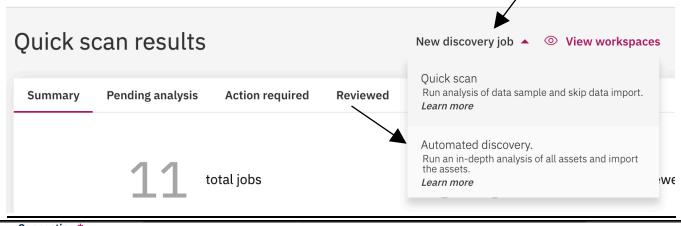
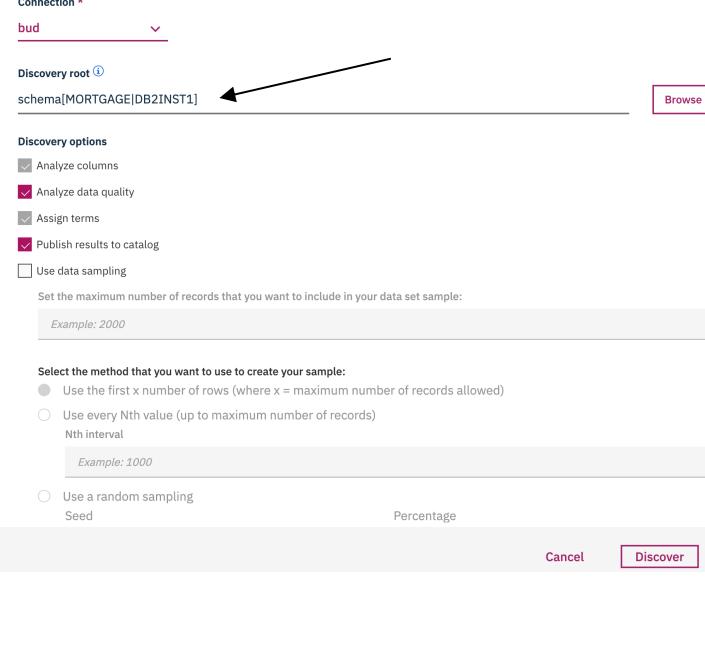
 <p>New governance rule</p> <p>Governance rule name Income cannot be null</p> <p>Description Income column must have a valid value</p> <p>Cancel Save as draft</p>	<p>On the New governance rule window create a rule with following information and click on Save as draft:</p> <p>Name: Income cannot be null Description: Income column must have a valid value</p> <p>It will take few minutes to appear under list of available rules. Once the new rule is available, publish it.</p>
--	--

 <p>Income cannot be null DRAFT Not started Delete draft Publish</p> <p>Overview Related content</p> <p>No secondary category added yet.</p> <p>Parent policies Add policy</p>	<p>Click on Add policy under Parent policies to assign the Data Validation rule to it.</p>
---	---

7.5. Automated Discovery

Re-run discover assets to add data to the catalog. During the discovery the data is imported, analyzed, and classified according the glossary you imported/created earlier.

	<ul style="list-style-type: none"> <input type="checkbox"/> Projects <input type="checkbox"/> Connections <input type="checkbox"/> My instances <input type="checkbox"/> Collect <input checked="" type="checkbox"/> Organize All catalogs Information assets Data and AI governance Curation Metadata import <p style="margin-top: 10px;">Data discovery</p>	<p>From Organize option on the left pane, choose Curation > Data discovery.</p>
--	---	--

	<p>To automated discover job</p> <p>Navigate to New discover job > Automated discovery</p>
	<p>To discover assets</p> <p>Choose the connection named bud that you created previously.</p> <p>Select Discover root as MORTGAGE > DB2INST1</p> <p>Check necessary Discover options</p> <p>Select Workspace as Mortgage.</p> <p>Click on Discover</p> <p>Wait till import and analyze phase complete.</p>

7.6. Add rule to metadata

Go to **Organize > Information assets**

Search for Database Table name **MORTGAGE_CUSTOMER**

Explore information assets

Database Table MORTGAGE_CUSTOMER

MORTGAGE_CUSTOMER
jdbc:db2://52.117.27.156:50000/MORTGAGE > MORTGAGE > DB2INST1

MORTGAGE_CUSTOMER
jdbc:db2://52.117.27.156:50000/MORTGAGE > MORTGAGE > DB2INST1

Click on **MORTGAGE_CUSTOMER** data set

Workspaces Catalog data sets

Filter results Data sets

Search data set

Schema Host name Created by Created on Modified by

Data sets 7 results

+ Add to workspaces

Name	Quality score	Threshold	First imported	Last published	Terms	Workspaces
MORTGAGE_CUSTOMER	99%	80%	Dec 2, 2019, 2:34 PM	Dec 2, 2019, 2:37 PM	12	1
MORTGAGE_CUSTOMER	99%	80%	Dec 2, 2019, 12:14 PM	Dec 2, 2019, 12:14 PM	1	

Database Table details
MORTGAGE_CUSTOMER

Governance Context: jdbc:db2://10.208.125.125:50000/MORTGAGE db2 > DB2INST1

Database Columns (10)

Created by admin.admin
Created on 04 June 2019, 11:28:49 am
Modified by InformationServerSystemUser
Modified on 04 June 2019, 11:28:49 am

Database Columns

- APPLIED_ONLINE
jdbc:db2://10.208.125.125:50000/MORTGAGE > db2 > DB2INST1 > MORTGAGE_CUSTOMER
- CARD_DEBT
jdbc:db2://10.208.125.125:50000/MORTGAGE > db2 > DB2INST1 > MORTGAGE_CUSTOMER
- CURRENT_LOANS
jdbc:db2://10.208.125.125:50000/MORTGAGE > db2 > DB2INST1 > MORTGAGE_CUSTOMER
- ID
jdbc:db2://10.208.125.125:50000/MORTGAGE > db2 > DB2INST1 > MORTGAGE_CUSTOMER
- INCOME**
jdbc:db2://10.208.125.125:50000/MORTGAGE > db2 > DB2INST1 > MORTGAGE_CUSTOMER

On Database Table Details window choose **Database Columns** from left

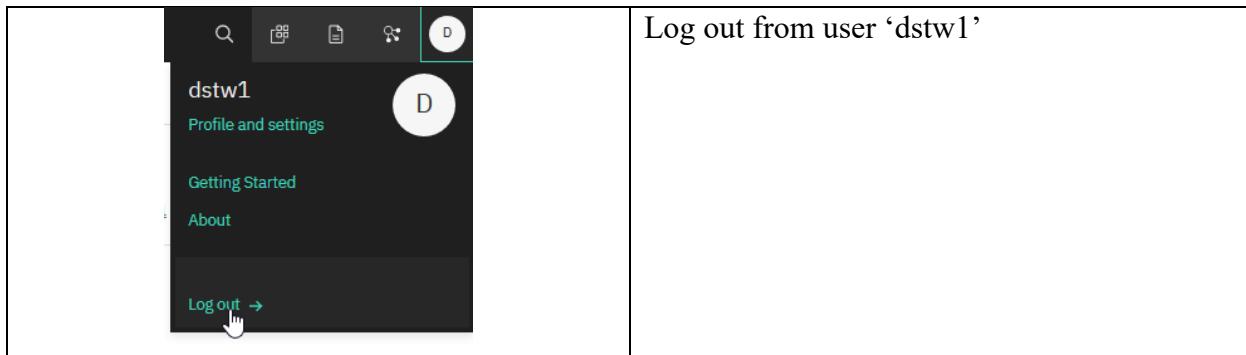
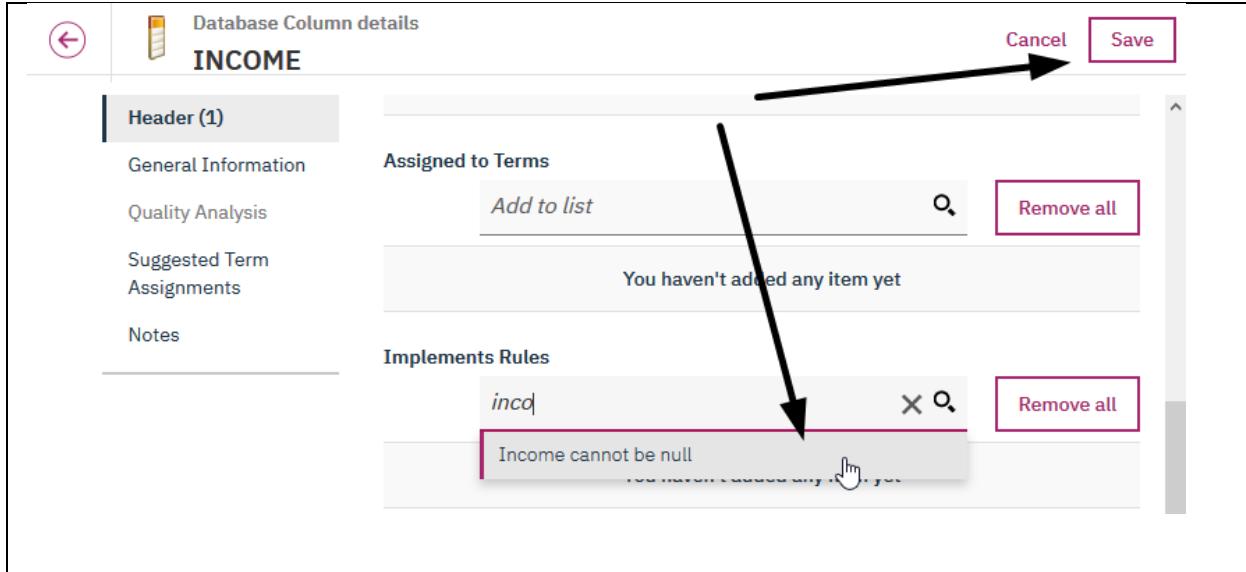
Select INCOME column

Next click on icon (right top corner) and choose Edit

Scroll down to **Implement Rules** section

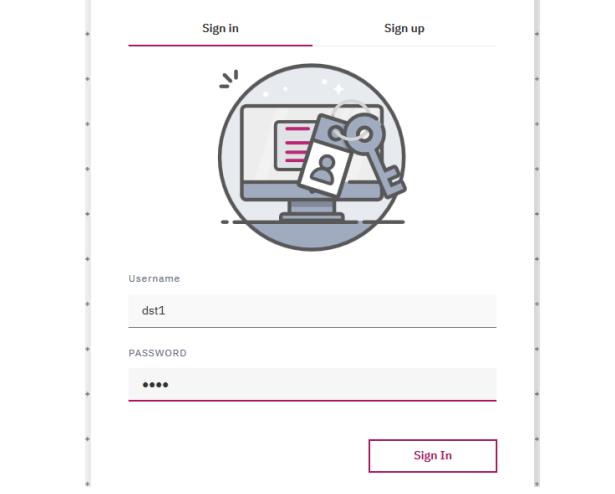
Search and select the rule **Income cannot be null** that you created earlier.

Click on **Save**

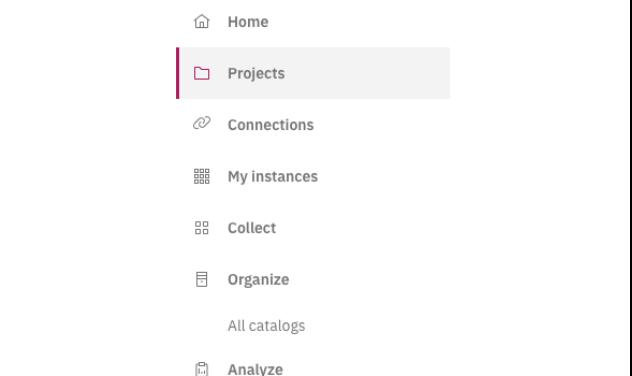


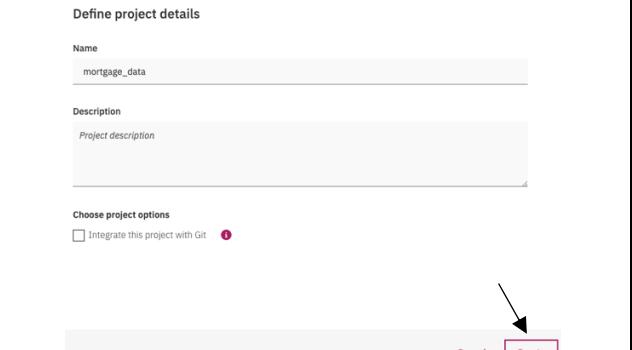
8. Access data as a Data Scientist

Explore the data require for build a model

	<p>Sigh in to the Cloud Pak for Data web console as user ‘dst1’ and password is ‘dst1’ that you created earlier.</p>
---	--

8.1. Create analytic project

	<p>Create a new analytical project by ‘Projects’ from right pane.</p> <p>Click on the  New project icon</p> <p>Select Create an empty project</p>
--	--

	<p>Provide project name ‘mortgage_data’ and click Create</p>
---	---

8.2. Assets from Glossary

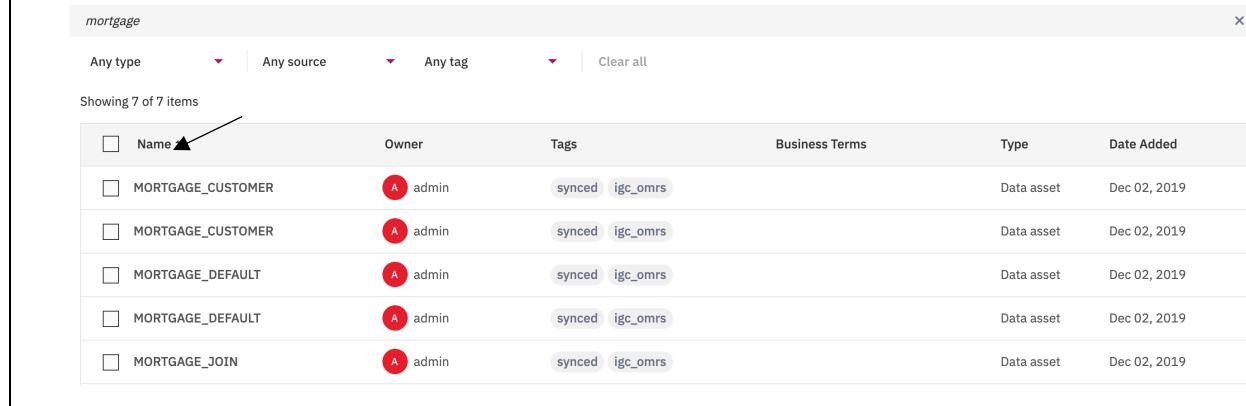
Let's look for mortgage related terms in glossary to get an idea about different data assets available on the system.

Go to **Organize > All catalogs** and choose **Default Catalog**

Search for word **Mortgage** from **Browse Assets** to find all mortgage related assets.

Click on each assets for additional information.

Default Catalog

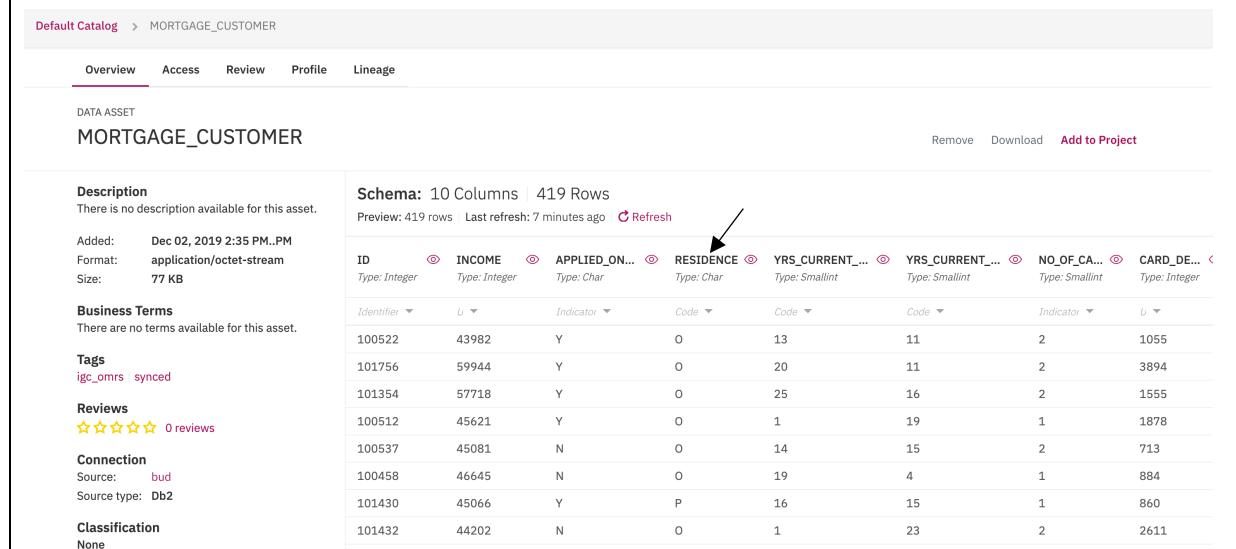


Name	Owner	Tags	Business Terms	Type	Date Added
MORTGAGE_CUSTOMER	admin	synced igc_omrs		Data asset	Dec 02, 2019
MORTGAGE_CUSTOMER	admin	synced igc_omrs		Data asset	Dec 02, 2019
MORTGAGE_DEFAULT	admin	synced igc_omrs		Data asset	Dec 02, 2019
MORTGAGE_DEFAULT	admin	synced igc_omrs		Data asset	Dec 02, 2019
MORTGAGE_JOIN	admin	synced igc_omrs		Data asset	Dec 02, 2019

8.3. Check Asset Details

Go through each data assets related to mortgage in glossary to have better idea about data you need for your project. For example, check the MORTGAGE_CUSTOMER.

The asset **Overview** tab shows the asset properties, such as the description, tags, format, size, and date added. You'll see a preview of the contents of the asset if the asset type supports previews and you have the proper permissions. Check individual column header description.



MORTGAGE_CUSTOMER									
DATA ASSET									
Overview									
Description There is no description available for this asset.									
Added: Dec 02, 2019 2:35 PM..PM Format: application/octet-stream Size: 77 KB									
Business Terms There are no terms available for this asset.									
Tags igc_omrs synced									
Reviews ☆☆☆☆☆ 0 reviews									
Connection Source: bud Source type: Db2									
Classification None									
Schema: 10 Columns 419 Rows Preview: 419 rows Last refresh: 7 minutes ago Refresh									
ID Type: Integer	INCOME Type: Integer	APPLIED_ON... Type: Char	RESIDENCE Type: Char	YRS_CURRENT_... Type: Smallint	YRS_CURRENT_... Type: Smallint	NO_OF_CA... Type: Smallint	CARD_DE... Type: Integer		
Identifier	U	Indicator	Code	Code	Code	Indicator	U		
100522	43982	Y	0	13	11	2	1055		
101756	59944	Y	0	20	11	2	3894		
101354	57718	Y	0	25	16	2	1555		
100512	45621	Y	0	1	19	1	1878		
100537	45081	N	0	14	15	2	713		
100458	46645	N	0	19	4	1	884		
101430	45066	Y	P	16	15	1	860		
101432	44202	N	0	1	23	2	2611		

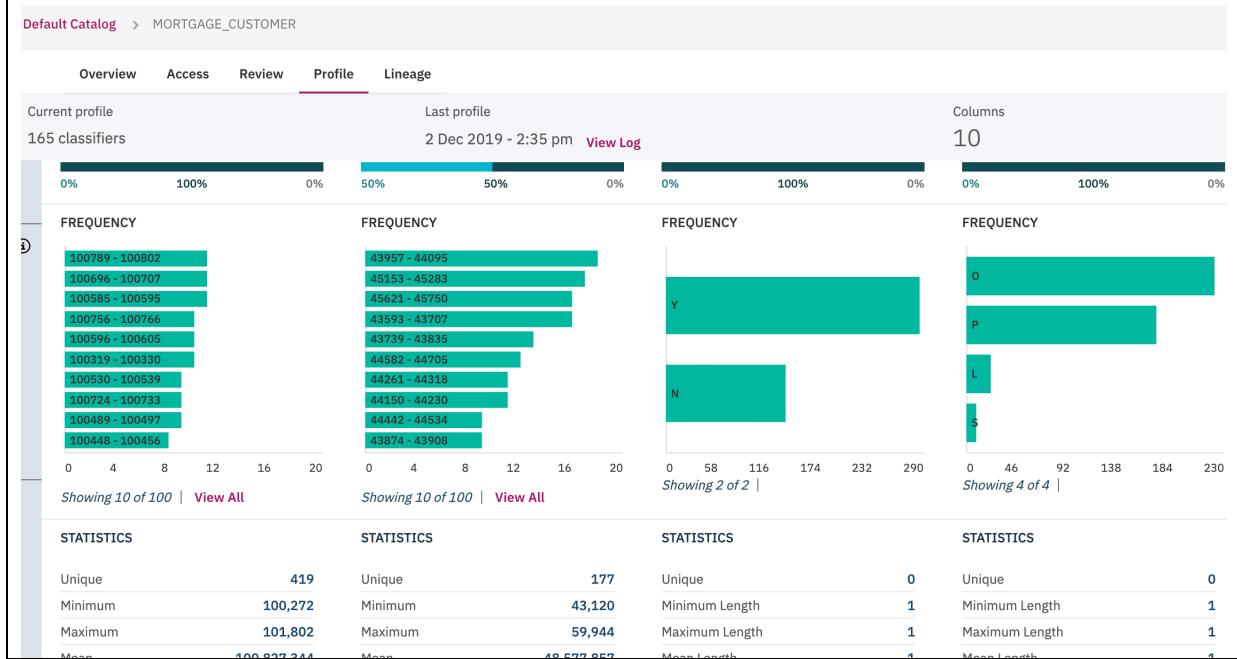
The **Review** tab shows the ratings and reviews of the asset by catalog collaborators. You can rate the asset and write a review on this page.

Overall Rating
0.0
★ ★ ★ ★ 0 reviews

Review Summary
5 (0)
4 (0)

My Review
dst1 | Dec 02, 2019
★ ★ ★ ★
Write a review of this asset to help others.

The **Profile** tab shows profile information about the contents of the asset. The profile of a data asset includes generated metadata and statistics about the textual content of the data. It contains relational or structured data shows information about each column in the data set, based on the first 5000 rows of data. The profile shows the frequency of the inferred data classes and statistics about the data for each column.



Target*
mortgage_data

Selected assets (1)

Asset Name	Catalog	Connection
MORTGAGE_CUSTOMER	Default Catalog	bud

Add

Once find right data asset use **Add to Project** tab to include it in your project. Select **Target** project as ‘mortgage_data’ and click on **Add**.

Similar way add MORTGAGE_DEFAULT and MORTGAGE_PROPERTY assets to your project.

9. Data Virtualization

Many time as a data engineer, you can receive requests for data from others. If you decide that a request requires data to be virtualized, You can use Data Virtualization (DV).

Assume you are a data engineer and need to deliver a data request that combined data sets of MORTGAGE_CUSTOMER, MORTGAGE_PROPERTY and MORTGAGE_DEFAULT.

DV allows integrate data sources across multiple types and locations and turns it into one logical data view. In this case, you have data across three different tables. Creating a virtual table you can quickly view data from different tables.

	<p>Sigh into the Cloud Pak for Data web console as user ‘deng1’ and password is ‘deng1’ that you created earlier.</p>
---	---

9.1. Adding a new data source for Db2

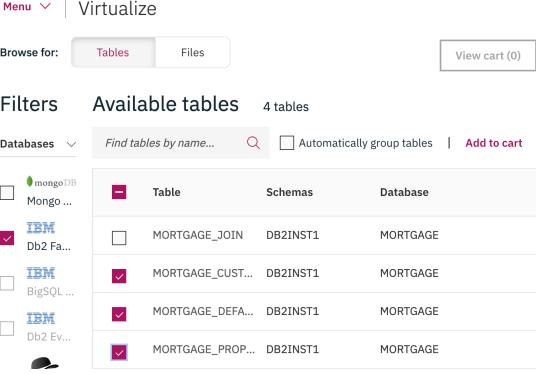
DV supports many relational and non-relational data sources (as well as files that reside on a local disk or network file system) that you can add to your data source ecosystem. After a data source has been added, any user that has virtualize permission can create virtual tables. DV agents connect to relational data sources using JDBC protocol. In this tutorial you will add a data source for Db2 database.

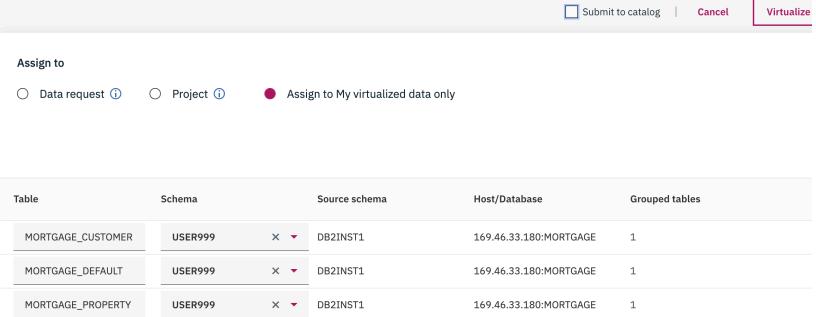
Define a data connection to Db2. Use your existing Db2 database connection for Db2 data source.

1. Go to **Collect > Data Virtualization > Menu > Data sources**
2. Click **Add > Add data source > Add connection**
3. Select **db2** that you created earlier and click **Next**

9.2. Select tables for virtualization

The most common mechanism for virtualizing data is to create a "view" or virtual table. Virtual tables can be full or segment of data from one or more tables. You can then run queries against the resulting virtual table.

<ul style="list-style-type: none"> Click Collect > Data virtualization > Menu > Virtualize Select tables MORTGAGE_CUSTOMER, MORTGAGE_PROPERTY and MORTGAGE_DEFAULT from MORTGAGE database, then click Add to cart Click View cart Click Next 	 <p>Virtualize</p> <p>Browse for: <input checked="" type="radio"/> Tables <input type="radio"/> Files View cart (0)</p> <p>Filters Available tables 4 tables</p> <p>Databases <input type="radio"/> mongoDB <input checked="" type="radio"/> IBM <input type="radio"/> BigSQL <input type="radio"/> IBM Db2 Event</p> <p>Find tables by name... <input type="checkbox"/> Automatically group tables <input type="button" value="Add to cart"/></p> <table border="1"> <thead> <tr> <th>Table</th> <th>Schemas</th> <th>Database</th> </tr> </thead> <tbody> <tr> <td>MORTGAGE_JOIN</td> <td>DB2INST1</td> <td>MORTGAGE</td> </tr> <tr> <td>MORTGAGE_CUST...</td> <td>DB2INST1</td> <td>MORTGAGE</td> </tr> <tr> <td>MORTGAGE_DEFA...</td> <td>DB2INST1</td> <td>MORTGAGE</td> </tr> <tr> <td>MORTGAGE_PROP...</td> <td>DB2INST1</td> <td>MORTGAGE</td> </tr> </tbody> </table>	Table	Schemas	Database	MORTGAGE_JOIN	DB2INST1	MORTGAGE	MORTGAGE_CUST...	DB2INST1	MORTGAGE	MORTGAGE_DEFA...	DB2INST1	MORTGAGE	MORTGAGE_PROP...	DB2INST1	MORTGAGE
Table	Schemas	Database														
MORTGAGE_JOIN	DB2INST1	MORTGAGE														
MORTGAGE_CUST...	DB2INST1	MORTGAGE														
MORTGAGE_DEFA...	DB2INST1	MORTGAGE														
MORTGAGE_PROP...	DB2INST1	MORTGAGE														

<ul style="list-style-type: none"> Uncheck the box for Submit to catalog Click Virtualize to complete the process 	 <p>Assign to</p> <p><input type="radio"/> Data request <small>(i)</small> <input type="radio"/> Project <small>(i)</small> <input checked="" type="radio"/> Assign to My virtualized data only</p> <p>Submit to catalog Cancel <input type="button" value="Virtualize"/></p> <table border="1"> <thead> <tr> <th>Table</th> <th>Schema</th> <th>Source schema</th> <th>Host/Database</th> <th>Grouped tables</th> </tr> </thead> <tbody> <tr> <td>MORTGAGE_CUSTOMER</td> <td>USER999</td> <td>X ▾</td> <td>DB2INST1</td> <td>169.46.33.180:MORTGAGE 1</td> </tr> <tr> <td>MORTGAGE_DEFAULT</td> <td>USER999</td> <td>X ▾</td> <td>DB2INST1</td> <td>169.46.33.180:MORTGAGE 1</td> </tr> <tr> <td>MORTGAGE_PROPERTY</td> <td>USER999</td> <td>X ▾</td> <td>DB2INST1</td> <td>169.46.33.180:MORTGAGE 1</td> </tr> </tbody> </table>	Table	Schema	Source schema	Host/Database	Grouped tables	MORTGAGE_CUSTOMER	USER999	X ▾	DB2INST1	169.46.33.180:MORTGAGE 1	MORTGAGE_DEFAULT	USER999	X ▾	DB2INST1	169.46.33.180:MORTGAGE 1	MORTGAGE_PROPERTY	USER999	X ▾	DB2INST1	169.46.33.180:MORTGAGE 1
Table	Schema	Source schema	Host/Database	Grouped tables																	
MORTGAGE_CUSTOMER	USER999	X ▾	DB2INST1	169.46.33.180:MORTGAGE 1																	
MORTGAGE_DEFAULT	USER999	X ▾	DB2INST1	169.46.33.180:MORTGAGE 1																	
MORTGAGE_PROPERTY	USER999	X ▾	DB2INST1	169.46.33.180:MORTGAGE 1																	

9.3. Creating virtual table

You can create a new virtual table based on existing tables under **My data** section. You can use "drag and drop" or write your own SQL to create the view.

<ul style="list-style-type: none"> Click Collect > Data virtualization > Menu > SQL editor to access the editor. Copy the following SQL statement and paste it on the editor Click on Run all 	<pre>CREATE VIEW MORTGAGE_JOIN_VIEW AS SELECT A.ID, INCOME, APPLIED_ONLINE, RESIDENCE, YRS_CURRENT_ADD, YRS_CURRENT_EMP, NO_OF_CARDS, CARD_DEBT, CURRENT_LOANS, LOAN_AMOUNT, SALE_PRICE, LOCATION, MORTGAGE_DEFAULT FROM MORTGAGE_CUSTOMER A, MORTGAGE_PROPERTY B, MORTGAGE_DEFAULT C WHERE A.ID = B.ID AND A.ID = C.ID;</pre>
---	--

Menu ▾ | SQL editor

* Untitled - 1

Syntax assistant | ⚙️

```

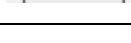
1
2
3 CREATE VIEW MORTGAGE_JOIN_VIEW
4 AS
5   SELECT A.ID, INCOME, APPLIED_ONLINE, RESIDENCE, YRS_CURRENT_
6     YRS_CURRENT_EMP, NO_OF_CARDS, CARD_DEBT, CURRENT_LOAN
7     LOAN_AMOUNT, SALE_PRICE, LOCATION, MORTGAGE_DEFAULT
8   FROM MORTGAGE_CUSTOMER A,
9     MORTGAGE_PROPERTY B,
10    MORTGAGE_DEFAULT C
11  WHERE A.ID = B.ID
12  AND A.ID = C.ID;
13

```

- Click **Collect > Data virtualization > Menu > My virtualized data** to access the virtual table MORTGAGE_JOIN_VIEW
- Check the box associated with MORTGAGE_JOIN_VIEW
- Click on the table actions menu 
- Select **Manage access** option
- On grant access window select All data virtualization users
- Click **Add**

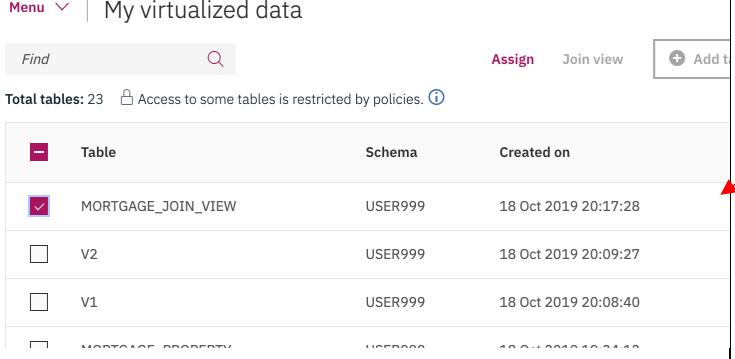
Grant access to

All data virtualization users  Specific users 

Users	Roles
<input style="width: 100%; height: 20px; border: 1px solid #ccc;" type="text" value="Search"/>   	
	
	

9.4. Add virtual table to catalog

Once you create a virtual table, you can add it to the catalog, making it easily searchable.

<ul style="list-style-type: none"> • Click Collect > Data virtualization > Menu > My virtualized data to find the virtual table just created. • Mark the checkbox associated with virtual table • Choose Submit to catalog from table action • Click on Confirm 	 <p>Menu My virtualized data</p> <p>Total tables: 23 Access to some tables is restricted by policies. ⓘ</p> <table border="1"> <thead> <tr> <th>Table</th> <th>Schema</th> <th>Created on</th> </tr> </thead> <tbody> <tr> <td><input checked="" type="checkbox"/> MORTGAGE_JOIN_VIEW</td> <td>USER999</td> <td>18 Oct 2019 20:17:28</td> </tr> <tr> <td><input type="checkbox"/> V2</td> <td>USER999</td> <td>18 Oct 2019 20:09:27</td> </tr> <tr> <td><input type="checkbox"/> V1</td> <td>USER999</td> <td>18 Oct 2019 20:08:40</td> </tr> <tr> <td><input type="checkbox"/> MORTGAGE_PROPERTY</td> <td>USER999</td> <td>18 Oct 2019 20:08:40</td> </tr> </tbody> </table>	Table	Schema	Created on	<input checked="" type="checkbox"/> MORTGAGE_JOIN_VIEW	USER999	18 Oct 2019 20:17:28	<input type="checkbox"/> V2	USER999	18 Oct 2019 20:09:27	<input type="checkbox"/> V1	USER999	18 Oct 2019 20:08:40	<input type="checkbox"/> MORTGAGE_PROPERTY	USER999	18 Oct 2019 20:08:40
Table	Schema	Created on														
<input checked="" type="checkbox"/> MORTGAGE_JOIN_VIEW	USER999	18 Oct 2019 20:17:28														
<input type="checkbox"/> V2	USER999	18 Oct 2019 20:09:27														
<input type="checkbox"/> V1	USER999	18 Oct 2019 20:08:40														
<input type="checkbox"/> MORTGAGE_PROPERTY	USER999	18 Oct 2019 20:08:40														

9.5. Publish virtualized table

A data steward needs approve the published request before the asset is added to the global data catalog. You signed in as user ‘admin’, it should allow to publish the virtual table.

 <p>Pending Publish to Catalog Requests</p> <table border="1"> <thead> <tr> <th>Name</th> <th>Type</th> <th>Project</th> <th>Owner</th> <th>Date Updated</th> <th>Status</th> </tr> </thead> <tbody> <tr> <td>> USER999.MORTGAGE_JOIN_VIEW</td> <td>view</td> <td>-</td> <td>admin</td> <td>21 October 2019, 2:49PM</td> <td>Pending</td> </tr> <tr> <td>> USER999.Currency USER999.Country</td> <td>table</td> <td>-</td> <td>admin</td> <td>17 October 2019, 8:40AM</td> <td>Pending</td> </tr> </tbody> </table>	Name	Type	Project	Owner	Date Updated	Status	> USER999.MORTGAGE_JOIN_VIEW	view	-	admin	21 October 2019, 2:49PM	Pending	> USER999.Currency USER999.Country	table	-	admin	17 October 2019, 8:40AM	Pending	<ul style="list-style-type: none"> • Click on  access the Home page • Click on Pending Publish to Catalog Requests • Click on  icon on left for virtual table MORTGAGE_JOIN_VIEW that you created • Click on Approve
Name	Type	Project	Owner	Date Updated	Status														
> USER999.MORTGAGE_JOIN_VIEW	view	-	admin	21 October 2019, 2:49PM	Pending														
> USER999.Currency USER999.Country	table	-	admin	17 October 2019, 8:40AM	Pending														

9.6. Access information for virtual table

To access virtual table from external application, you need the JDBC connection information. Click on **Collect > Data Virtualization > Menu > Service settings** to find out access information. You will use this information later in the building model section.

The screenshot shows the 'Access information' page under 'Add-on settings'. It displays the following details:

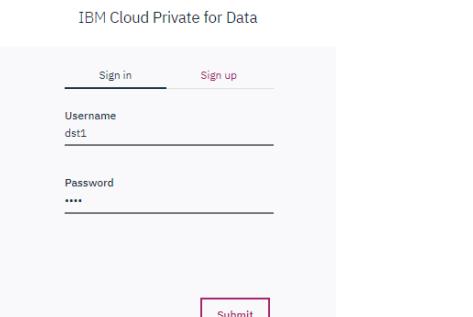
User ID	user999	Copy
Password Show	Copy Refresh
JDBC connection URL	jdbc:db2://dv-server.zen.svc.cluster.local:32051/...	Copy

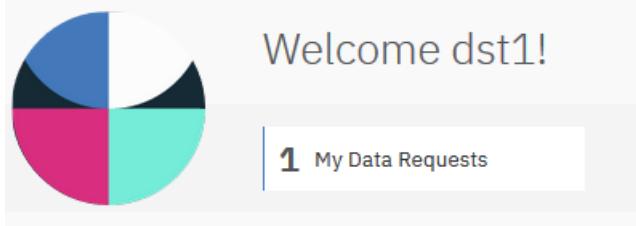
The screenshot shows the IBM Cloud Private for Data dashboard. A red arrow points to the user profile icon in the top right corner. A dropdown menu is open, showing 'Signed in as: deng1', 'Getting Started', 'Settings', and 'Sign Out'. The 'Sign Out' button is highlighted with a red box.

Sign out from user **deng1**

10. Build Model

With Cloud Pak for Data, you can collaborate with other team members on analytic projects to create visualizations and machine learning models with data from your enterprise. In this step you will build a simple model to predict the possibilities of mortgage default by customer. The object of this model is to show the functionality of Cloud Pak for Data, not the prediction accuracy. One can use lot more data and build a complex algorithm to get better accuracy.

	<p>Sign: in to the Cloud Pak for Data web console as user ‘dst1’ and password is ‘dst1’ that you created earlier.</p>
---	---

	<p>At this point data engineer deliver the data set for the data you requested. You can go to the home page by clicking on  icon from left pane</p>
--	---

10.1. Navigate to analytics project

Select **Projects** option from the left pane and click on the analytics project ‘mortgage_data’ that you created earlier.

10.2. Create deployment space

Create a separate deployment space for your project ‘mortgage_data’.

<p>Choose : My Projects > mortgage_data > Settings > Associate a deployment space > New</p> <p>Connect to a deployment space</p> <p>New Existing</p>		<p>Name new deployment space as ‘MortgageDeploymntSpace’</p> <p>Click on Associate</p>
<p>Name</p> <p>MortgageDeploymentSpace</p>		
<p>Description (Optional)</p> <p>Description of deployment space</p>		
<p>Cancel Associate</p>		

10.3. Create notebook

Create a notebook from a predefined Jupyter notebook that available on Github.

- Go to : My Projects > **mortgage_data** > Add to project
- Chose asset type as Notebook
- The new notebook needs to create from URL
- Name the notebook as **MortgageNotebook**
- Use notebook URL as <https://github.com/IBM-ICP4D/icp4d-tutorials/blob/master/assets/mortgage-002/MortgageNotebook.V25.jupyter-py36.ipynb>
- Click on **Create Notebook**

My Projects > mortgage_data > Add Notebook

New notebook

Blank From file **From URL**

Name
MortgageNotebook 24 characters remaining

Description (optional)
Type your Description here 500 characters remaining

Select runtime
Default Python 3.6 (1 vCPU and 2 GB RAM)

Notebook URL
<https://github.com/IBM-ICP4D/icp4d-tutorials/blob/master/assets/mortgage-002/MortgageNotebook.V25>

Create Notebook

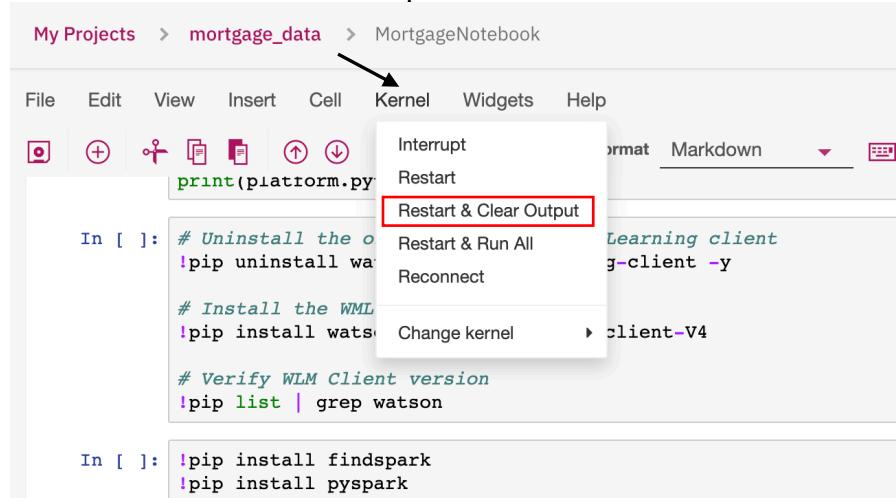
10.4. Review and run notebook

The majority of the code in the notebook is standard open source code that's used for various steps in the predictive analytics process.

Switch to edit mode by clicking on icon from top of the screen.

Do not run all cells at once. Follow the instruction below to run the notebook.

Run the **Step 1: Intall** section first. Once all package installed make sure restart the Python kernel before move on next step.



The screenshot shows a Jupyter Notebook interface with the following details:

- Path:** My Projects > mortgage_data > MortgageNotebook
- Kernel Menu:** The 'Kernel' menu is open, showing options: Interrupt, Restart, and a submenu with 'Restart & Clear Output' (which is highlighted with a red box), 'Restart & Run All', 'Reconnect', and 'Change kernel'.
- Code Cells:**
 - In []:** print(platform.py)
 - In []:** # Uninstall the old Watson API

```
!pip uninstall watson
```
 - In []:** # Install the WML

```
!pip install wats
```
 - In []:** # Verify WLM Client version

```
!pip list | grep watson
```
 - In []:** !pip install findspark

```
!pip install pyspark
```

Action: restart the kernel!

Go the **Step 2: Authenticate** section and update the **url**, **username** and **password** fields with your CPD UI console details and access credential.

Step 2: Authenticate

```
[ ]: WML_CREDENTIALS = {
    "instance_id": "openshift",
    "url" : "https://zen-cpd-zen.apps.testcluster.demo.ibmcloud.com",
    "username": "admin",
    "password": "passw0rd",
    "version": "2.5.0"
}
```

In the next notebook cell, update the **dsn_url**, **dsn_uid** and **dsn_pwd** values with the information available from **Collect > Virtualized data > Menu > Add-on settings**.

```
[ ]: #Enter the values for you database connection found under data virtualization
dsn_url = "jdbc:db2://dv-server.zen.svc.cluster.local:32051/bigsql" # e.g.
dsn_uid = "user1022" # e.g.
dsn_pwd = "sw?#@lt_674MfPI5" # e.g.
```

Run all cells between step 2 and 6.

You may need to change the MORTGAGE_JOIN_VIEW schema name in step 3, according to your environment.

On **Step 7: Set default space**, run the first cell and find out the **GUID** for space name **MortgageDeploymentSpace**.

On the next cell replaced the **GUID** with one that you found above.

```
In [ ]: # Example: client.set.default_space('b49e13e8-ec68-408d-84a1-957e28c154b1')
client.set.default_space('GUID')
```

Run through remaing cells, so that it generates and deployed the model.

Before exit, save the notebook .

12.5. Test the model

Go to: Analyze > Analytics deployment to access deployed model

Select the **MortgageDeploymentSpace** from the list of analytic deployment space

Click on the **MORTGAGE PREDICTION MODEL**

Choose the **MORTGAGE PREDICTION** model

Click on **Test** tab

The screenshot shows the 'Analytics deployment spaces' interface. The path is: Analytics deployment spaces > MortgageDeploymentSpace > MORTGAGE PREDICTION MODEL > MORTGAGE PREDICTION. The 'Test' tab is selected. On the left, there's a 'Body' section with a placeholder 'Paste the request payload here' and a 'Predict' button. On the right, there's a 'Result' section which is currently empty. To the right of the result section, there's a sidebar with the following details:

- MORTGAGE PREDICTION** (highlighted with a red circle)
- Deployed** (green checkmark)
- Created**: Nov 07, 2019 11:48 PM
- Updated**: Nov 08, 2019 06:21 PM
- Deployment ID**: b7a58231-fd99-4d9f-a760-7d81... (with a copy icon)
- Software**: /v4/runtimes/spark-mllib_2.3
- Description**: No description provided
- Associated asset**: MODEL MORTGAGE PREDICTION M...
- Model ID**: 4809b65e-9cab-4870-b93c-7444... (with a copy icon)

```
{
  "input_data": [
    {
      "fields": [
        "INCOME",
        "APPLIED_ONLINE",
        "RESIDENCE",
        "YRS_CURRENT_ADD",
        "YRS_CURRENT_EMP",
        "NO_OF_CARDS",
        "CARD_DEBT",
        "CURRENT_LOANS",
        "LOAN_AMOUNT",
        "SALE_PRICE",
        "LOCATION"
      ],
      "values": [
        [
          43151,
          "N",
          "P",
          6,
          9,
          1,
          750,
          1,
          8600,
          320000,
          110
        ]
      ]
    }
  ]
}
```

Copy this sample data and paste it on the **Enter input data** box.

Click on **Predict**

According on input values, model will predict and displays the result.

Analytics deployment spaces > MortgageDeploymentSpace > MORTGAGE PREDICTION MODEL > MORTGAGE PREDICTION

ONLINE
MORTGAGE PREDICTION

API reference **Test**

Enter input data

Body

```
{
  "input_data": [
    {
      "fields": [
        "INCOME",
        "APPLIED_ONLINE",
        "RESIDENCE",
        "YRS_CURRENT_ADD",
        "YRS_CURRENT_EMP",
        "NO_OF_CARDS"
      ]
    }
  ]
}
```

Predict

Result

```

0  {
1    "predictions": [
2      {
3        "fields": [
4          "INCOME",
5          "APPLIED_ONLINE",
6          "RESIDENCE",
7          "YRS_CURRENT_ADD",
8          "YRS_CURRENT_EMP",
9          "NO_OF_CARDS",
10         "CARD_DEBT",
11         "CURRENT_LOANS",
12         "LOAN_AMOUNT",
13         "SALE_PRICE",
14         "LOCATION",
15         "MORTGAGE_DEFAULT",
16         "AppliedOnlineEncoded"
      ]
    }
  ]
}
```

MORTGAGE PREDICTION
Deployed

Created
Nov 07, 2019 11:48 PM

Updated
Nov 08, 2019 06:32 PM

Deployment ID
b7a58231-fd99-4d9f-a760-7d81...

Software
/v4/runtimes/spark-mllib_2.3

Description
No description provided

Associated asset

MODEL
MORTGAGE PREDICTION M...

Model ID
4809b65e-9cab-4870-b93c-7444...