UDS registration runbook

Contains documentation on the different aspects of Registration and its component. This document shows a new or existing UDS 1.1 user about navigating in the UDS portal UI and how to work with different Registration components

UDS 1.1 Registration Draft

Author

| **Title** | **Name** | **Update date** |
| --- | --- | --- |
| Business Analyst/Scrum Master | Ameya Gaikwad | 09/14/2017 |

# Introduction

Unified Data Service (UDS) is a software suite of common data services that are modular and extensible. These services are designed for use in Big Data environments. UDS provides the several features and benefits including:

* Eliminates costs by providing the entire firm with a set of common generic data services such as data ingestion, registration, encryption, processing, user access etc. that can be implemented in multiple data environments.
* Accelerates progress by enabling JPMC to quickly bootstrap new data environments, start ingesting data and perform analytics/discovery of data.  Leverages software foundation
* Provides significant business value by providing a way to link disparate data environments into a common Virtual Data Lake so that management can “see data across the environments”.
* Manages metadata in a centralized metadata repository with linkage to the Logical Data Model and the Firm Wide Data Model (FDM)
* Architected as a collection of Micro services. User community will be both back end IT teams working on implementing data environments as well as front end Business consumers of data

# Contacts

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name** | **Role** | **Email** | **Phone** | **Location** |
| Ramesh Thumma | Product Owner | [ramesh.r.thumma@jpmchase.com](mailto:ramesh.r.thumma@jpmchase.com) | (212) 622-1945 | NY |
| Wang Dong | Dev Lead | [dong.x.wang@chase.com](mailto:dong.x.wang@chase.com) | (614) 217-9465 | OH |
| Ameya Gaikwad | Business Analyst/Scrum Master | [ameya.k.gaikwad@jpmchase.com](mailto:ameya.k.gaikwad@jpmchase.com) | (718) 242-8200 | NY |
| Santanu Nandi | Developer | [santanu.nandi@jpmchase.com](mailto:santanu.nandi@jpmchase.com) | (718-242-8180) | NY |
| Narendra Sripada | Developer | [narendra.k.sripada@chase.com](mailto:narendra.k.sripada@chase.com) |  | NY |
| Mikhail Grinshteyn | Developer | [mikhail.grinshteyn@jpmchase.com](mailto:mikhail.grinshteyn@jpmchase.com) |  | NY |
| Nanda Gajjela | Developer | [nanda.k.gajjela@jpmchase.com](mailto:nanda.k.gajjela@jpmchase.com) | (614) 248-4174 | OH |
| Thiagaraja Gurusamy | Developer | [thiagaraja.gurusamy@jpmchase.com](mailto:thiagaraja.gurusamy@jpmchase.com) | (614) 248-0791 | OH |
| Yanan Li | Developer | [yanan.li@chase.com](mailto:yanan.li@chase.com) | (614) 244-4708 | OH |
| Caitlin Van Gundy | Developer | [caitlin.vangundy@jpmorgan.com](mailto:caitlin.vangundy@jpmorgan.com) | (614) 248-4086 | OH |
| Manju Rao | QA | [manju.rao@jpmchase.com](mailto:manju.rao@jpmchase.com) |  | NY |
| Sajadul Hussain | QA | [sajadul.x.hussain@chase.com](mailto:sajadul.x.hussain@chase.com) |  | NJ |
| Sindhu Ganne | QA | [sindhu.ganne@chase.com](mailto:sindhu.ganne@chase.com) | (201) 595-1836 | NJ |
| Shobha Shivanna | QA | [shobha.shivanna@chase.com](mailto:shobha.shivanna@chase.com) |  | Bengaluru, India |

Table of Contents

[I. Introduction 2](#_Toc493153209)

[II. Contacts 3](#_Toc493153210)

[1. Data Set Groups 6](#_Toc493153211)

[1.1 Data Set Groups List Page 6](#_Toc493153212)

[1.2 Create a new Data Domain 8](#_Toc493153213)

[1.3 Edit a Data Set Group 10](#_Toc493153214)

[2. Data Set 10](#_Toc493153215)

[2.1 Data Set List Page 11](#_Toc493153216)

[2.2 Import a Data Set 14](#_Toc493153217)

[2.3 Create a Data Set 15](#_Toc493153218)

[2.3.1 Source type: Generic 17](#_Toc493153219)

[2.3.1.1 Source type: Generic and Registration type simple 18](#_Toc493153220)

[2.3.1.2 Source type: Generic and Registration type simple 20](#_Toc493153221)

[2.3.1.3 Source type: Generic and Registration type light 21](#_Toc493153222)

[2.3.2 Source type: File 21](#_Toc493153223)

[2.3.2.1 Registration type: Full 23](#_Toc493153224)

[2.3.2.2 Registration type: Light 25](#_Toc493153225)

[2.3.3 Source type: Table 26](#_Toc493153226)

[2.3.4 Source type: Directory 27](#_Toc493153227)

[2.4 Edit a Data Set 28](#_Toc493153228)

[2.4.1 Edit a Data Set l: file (Light) 28](#_Toc493153229)

[2.4.2 Edit a Data Set l: file (Full) 29](#_Toc493153230)

[2.4.3 Edit a Data Set l: Generic (Light) 31](#_Toc493153231)

[2.4.4 Edit a Data Set l: Generic (Simple) 31](#_Toc493153232)

[2.4.5 Edit a Data Set l: Generic (Complex) 32](#_Toc493153233)

[2.5 Delete, promote, Clone or add ETL Permissions to a Data Set 33](#_Toc493153234)

[3. Namespaces 36](#_Toc493153235)

[3.1 Namespace List Page 36](#_Toc493153236)

[3.2 Create a new Name space 37](#_Toc493153237)

[3.3 Edit a Namespace 37](#_Toc493153238)

[3.4 Delete a Namespace 38](#_Toc493153239)

[4. Encryption Policy 38](#_Toc493153240)

[4.1 View Encryption Policy 38](#_Toc493153241)

[4.2 Create an Encryption Policy 39](#_Toc493153242)

[4.3 Edit or Update an Encryption Policy 40](#_Toc493153243)

[4.4 Delete an Encryption Policy 41](#_Toc493153244)

[5. Processors 42](#_Toc493153245)

[5.1 Processor list page 43](#_Toc493153246)

[5.2 Create a new Processor 44](#_Toc493153247)

[5.3 Edit a Processor 45](#_Toc493153248)

[5.4 Delete a Processor 46](#_Toc493153249)

[6. Zones 47](#_Toc493153250)

[6.1 Zone list page 47](#_Toc493153251)

[6.2 Create a new Zone 48](#_Toc493153252)

[7. Data Flows 49](#_Toc493153253)

[7.1 Data Flow List Page 49](#_Toc493153254)

[7.2 Create a New Data Flow 50](#_Toc493153255)

[7.2.1 When Input ‘Option Types’= ‘Sql’ 53](#_Toc493153256)

[7.2.2 When Input ‘Option Types’= ‘File’ 54](#_Toc493153257)

[7.2.3 When Input ‘Option Types’= ‘Table’ 54](#_Toc493153258)

[7.2.4 When Output ‘Option Types’= ‘HDFS File’ 55](#_Toc493153259)

[7.2.5 When Output ‘Option Types’= ‘HDFS File’ 56](#_Toc493153260)

[7.3 Edit an existing Data Flow 57](#_Toc493153261)

[7.4 Delete an existing Data Flow 58](#_Toc493153262)

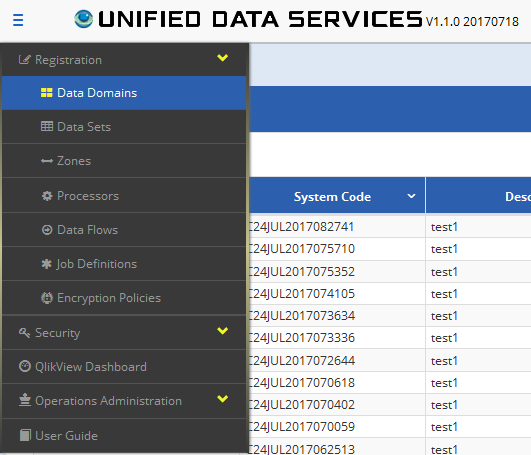
# Data Set Groups

A Data set group in UDS is a namespace for a particular set of data objects being transmitted to or from the Ecosystem. An example of a data domain is a System of Record that transmits one or many file structures to HDFS.

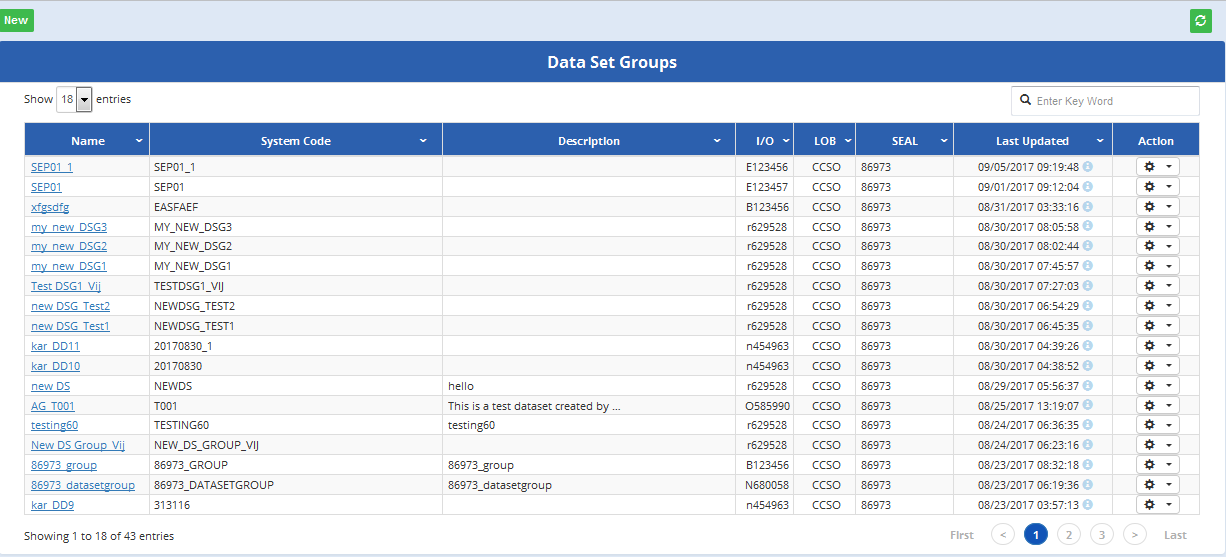
A data set group contains important information about the data domain including specification of the ownership of the data, Peregrine Queue, owning LOB and SEAL ID. These properties are utilized when servicing Data Domain requests like escalations, security requests etc.

# Data Set Groups List Page

To navigate to the data Domain list page, click on the menu bar on the top left



Data set group list Page displays all the Data Domains which have been created in the past. Below is the screenshot of the Data set group List page. Each column can be sorted in ascending or descending order by clicking on the  icon next to the selected column.



The Data set group list page shows the following fields:

|  |  |
| --- | --- |
| **Field Name** | **Description** |
| **Name** | This is the name of the Data Domain. Data Domain names are unique for each tenant |
| **System Code** | This is the system code for the Data Domain. |
| **Description** | A brief Description for the Data Domain. |
| **I/O** | Information Domain SID responsible from Data Domain. |
| **LOB** | Line of business which who this data domain has been created |
| **SEAL** | Seal ID of the tenant for which the data domain Originates. SEAL/UWAT ID for the application publishing SOR. |
| **Last Updated** | The last updated timestamp. |
| **Action** | A Data Domain cannot be deleted. Data Domain can perform the following actions: Promote |

The action column performs the following operations

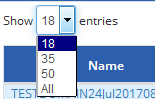


**Promote**: this promotes the data domain from lower environment to higher environment but it cannot promote from higher to the lower environment. For example it can promote from design to integration and/or integration to preview but it cannot other way around. For some Data domains, a user may not see a promote feature available. This is because they need to a have correct LDAP roles to promote. Below table shows what LDAP roles are required to the type of promotions.

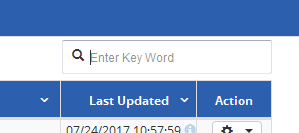
|  |  |  |
| --- | --- | --- |
| **Environment promotions** | **LDAP Role** | **LDAP Role Description** |
| From Design to Integration | UDS\_DEV\_META\_PROMOTER | Provide users ability to promote metadata for the following from Design to Preview: - data domain, datasets, encryption policies, data flows, data processors |
| From Integration to Preview | UDS\_QA\_META\_PROMOTER. | Provide users ability to promote metadata for the following from Integration to Preview: - data domain, datasets, encryption policies, data flows, data processors |

For a list of all the LDAP roles, please refer to <https://confluence.uk.jpmorgan.com/confluence/display/EDISON/LDAP+Roles>

By default, it will display the latest 18 data domains which have been created. But it can be customized by a “Show” entries drop down menu to view 18, 35, 50 or “All” data domains.



A user can search for data domain by typing Name, System Code, Description, I/O, LOB or Last Updated values in the search bar located at the top left of the list page.

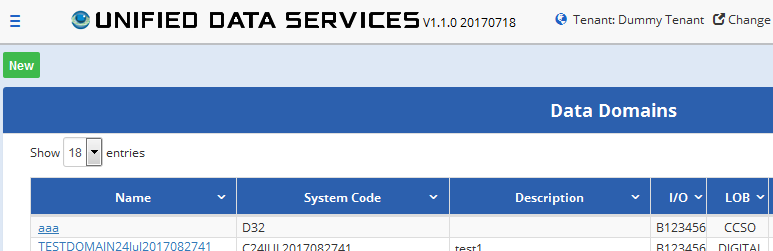


A user can navigate to a page on the list page by selecting the page in the bottom of the list Page. The list Page will also show what page a user is currently on.

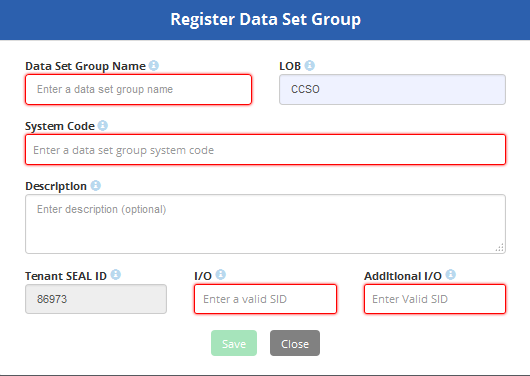


# Create a new Data Domain

To create a new Data Domain, click on the “New” green button on the top left of the list page.



A “Register Data Set Group” pop up page will appear to create a new data Domain



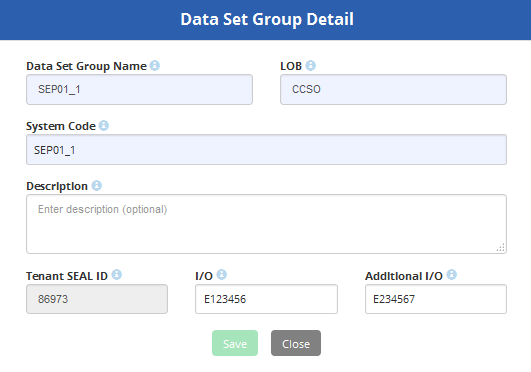
A user enters the information needed to create the Data Domain. The fields marked in red above are mandatory fields. Tenant Seal ID, I/O and Additional I/O are auto populated under which tenant the user has logged in as. If the entire mandatory to register a data domain are entered the “Save” button will be enabled. The save button will remain to be disabled until all the mandatory fields are entered. The grey “Close” button will be enabled at all times. By clicking on the “close” button, it will cancel the operation of Registering a new Data Domain and will take the user back on the Data Domain list page.

The field description and rules for registering a data set group is the below table:

|  |  |  |  |
| --- | --- | --- | --- |
| **Field Name** | **Description** | **Required** | **Data Type** |
| **Data Set Group Name** | This is the name of the Data Domain. Are Data Domain Names Unique?? | Y | VARCHAR2 (100 Bytes) |
| **LOB Owner** | Owning Line of Business of the Data Domain. The values are auto populated from the drop down menu. | Y | VARCHAR2 (10 Bytes) |
| **System Code** | System Code for the Data Domain | Y | VARCHAR2 (40 Bytes) |
| **Description** | A brief description of the Data Domain | N | VARCHAR2 (1000 Bytes) |
| **Tenant Seal ID** | Seal ID of the tenant for which the data domain Originates. This is auto populated. Default value = ' '. | Y | VARCHAR2 (30 Bytes) |
| **I/O** | Information Domain SID responsible from Data Domain. This is auto populated. | Y |  |
| **Additional I/O** | Additional Information owner SID responsible for Data Domain. This is auto populated. | Y |  |

# Edit a Data Set Group

Data Domains which have been created earlier can be edited. This is performed in the list page by clicking on the “Name”. After a user clicks on “Name” in the Data set group list page, a “Data Domain Detail” pop up appears as show in the below image where the user can update the values for already created data domain. Not all the fields are editable for an already created Data Domain. A user can edit the LOB owner and the description only.



# Data Set

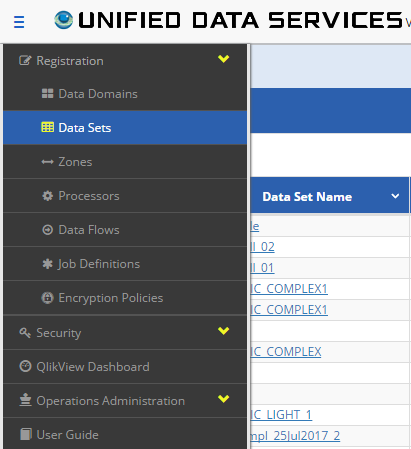
A dataset represents a source of data that resides/will reside in the Data Reservoir. A dataset specifies the properties of data including its source format, source and target locations, escalation resources, technical parsing properties and structure. This enables UDS to perform operations on data like parsing the incoming data (file/table), performing logical operations on it and escalating issues for it.

UDS currently supports the following types of dataset registrations:

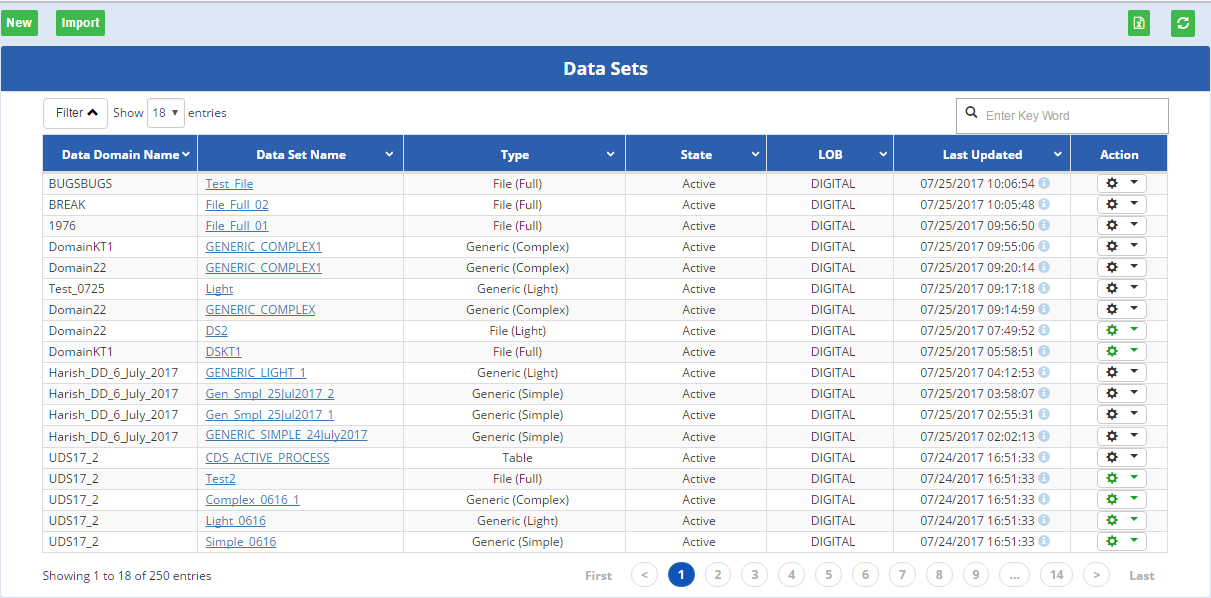
* File
  + Light
    - Used for easy archiving of files in the Data Reservoir
    - Files do not have a structure associated with them (no field registration required)
    - The primary importance for this type of registration is the ability to rapidly store and retrieve data in the ecosystem
    - An example use case is storing data for compliance reasons Used for complete registration of files and fields to be ingested into Data Reservoir
    - Utilizes the CDS data ingestion service to ingest a file based on a defined structure used to register structured datasets that need to go through ETL processing in the Ecosystem Used to register and ingest tables from a source database in the Data Reservoir
    - Based on the Source Resource details and the schema/dataset selected by user during registration, CDS imports the table structure from the source tables into the Target Zone Used to identify data that will be stored in a specific directory path in the target resource during ingestion
  + Full
    - Used for complete registration of files and fields to be ingested into Data Reservoir
    - Utilizes the CDS data ingestion service to ingest a file based on a defined structure
* Table
  + Used to register and ingest tables from a source database in the Data Reservoir
  + Based on the Source Resource details and the schema/dataset selected by user during registration, CDS imports the table structure from the source tables into the Target Zone
* Directory
  + Used to register and ingest data stored in a specific directory path in the Ecosystem
* Generic
  + Used to identify data that will be stored in a specific directory path in the target resource during ingestion

# Data Set List Page

To Navigate to the Data Set List Page from the main menu, Click “Registration” and select “Data Sets” from the drop down as show in the below image.

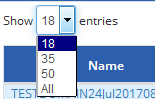


This will navigate you to the data set list page as shown below. The Data set list page displays the data sets which have been created. By default it is sort by the last updated value.

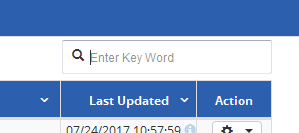


The list page shows the following fields. Each column can be sorted in ascending or descending order by clicking on the  icon next to the selected column.

By default, it will display the latest 18 data sets which have been created. But it can be customized by a “Show” entries drop down menu to view 18, 35, 50 or “All” data sets.



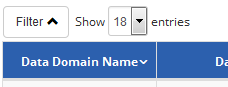
A user can search for data set by typing Data Domain Name, Data Set Name, Type, State, LOB or Last Updated values in the search bar located at the top left of the list page.



A user can navigate to a page on the list page by selecting the page in the bottom of the list Page. The list Page will also show what page a user is currently on.



A user can further filter to display specific data set information by clicking on the “Filter” button

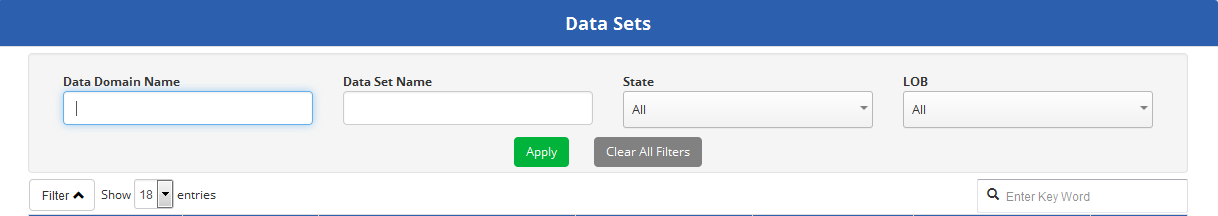


A user can filter by searching on the following fields as shown in the below image. Not all fields need to be entered to filter and search.

 Button- will execute the search based off on the filter condition

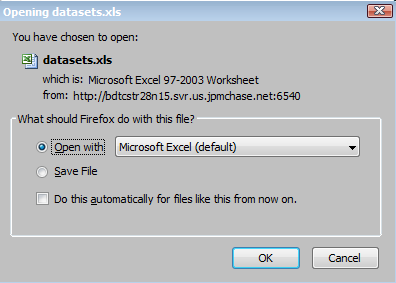
 Button- will clear the fields entered in the search fields.

If the Filter button is clicked again, the below search bar will disappear.



A user has an option to refresh the data set list page by clicking on the refresh  button the located on the top right of the list page.

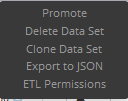
A user has an option to download the data sets by clicking on the  button located on the top right of the list page. By clicking on this, a pop up box will appear as shown in the below image asking the users permission to download or open the file. The file is downloaded in the xls format.



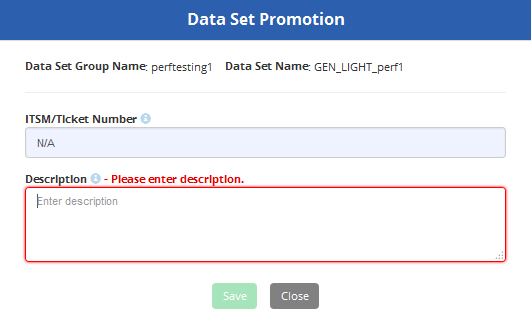


|  |  |
| --- | --- |
| **Field Name** | **Description** |
| **Data Domain Name** | The name of the data domain |
| **Data Set Name** | Data set with which is associated with |
| **Type** | The type of data domain  Simple  Generic (light)  Generic (complex)  Table  File (full) |
| **State** | Active - design, integration, preview, final  Pending Classification, Pending Approval - integration |
| **LOB** | The line of business |
| **Last Updated** | The date the last update took place |
| **Action** |  |

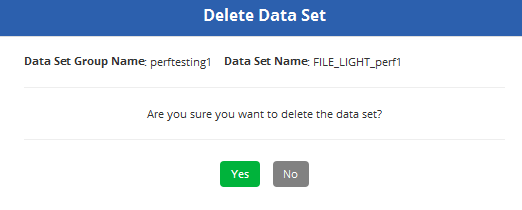
**Action**: following actions can be performed on a data set



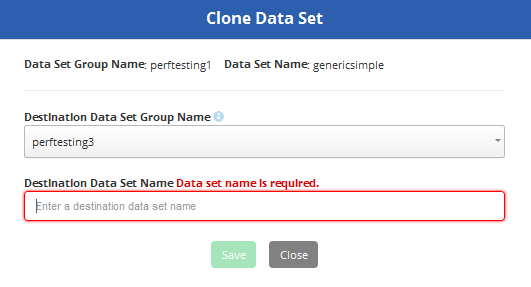
**Promote**: This is used to promote a data set from a lower to higher environment.



**Delete Data set:** This is used to delete a data set. A confirmation page will appear which will ask the user if he wants to proceed deleting the data set.



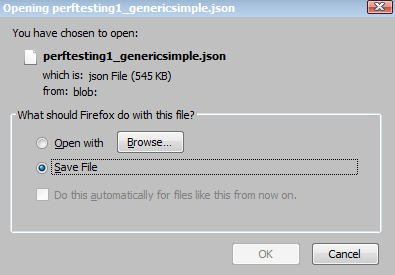
Clone data set. This is used to clone the data set. A destination data set group and destination data set name should be provided in order to clone the dataset.



Export to Jason: This is used to download the data set as a Jason file. A screen appears as shown below to confirm if the user wants to export the data set.

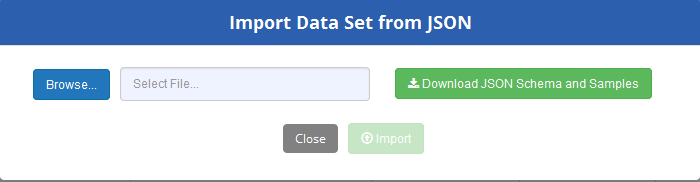


If the user clicks ‘yes’ the following screen appears to cobnfirm if the user wants to open the Json file or download and save the Json file.



# Import a Data Set

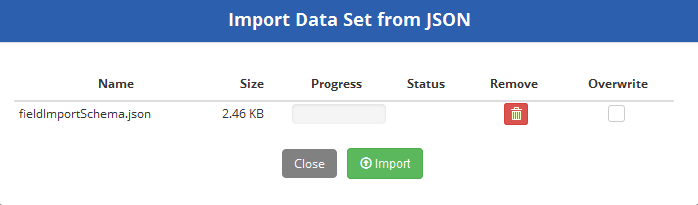
A new Data set can be imported by clicking on the  button located on top of the Data set list page. A new data set can be imported by uploading a new jSON File. After clicking on the import button, a new pop up for “Import Data Set from jSON appears as shown in the below image.



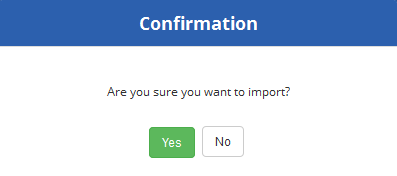
A sample of JSON Schema can be downloaded here



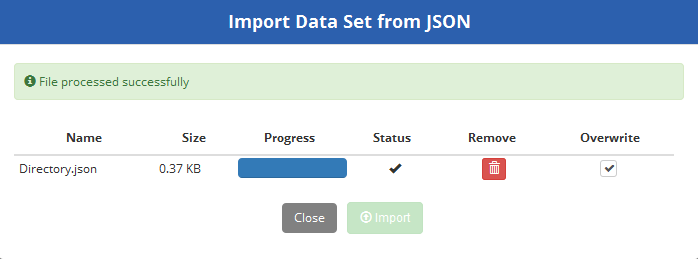
Browse for the jSON file want to upload and click import.



Confirm if you wish to import or not



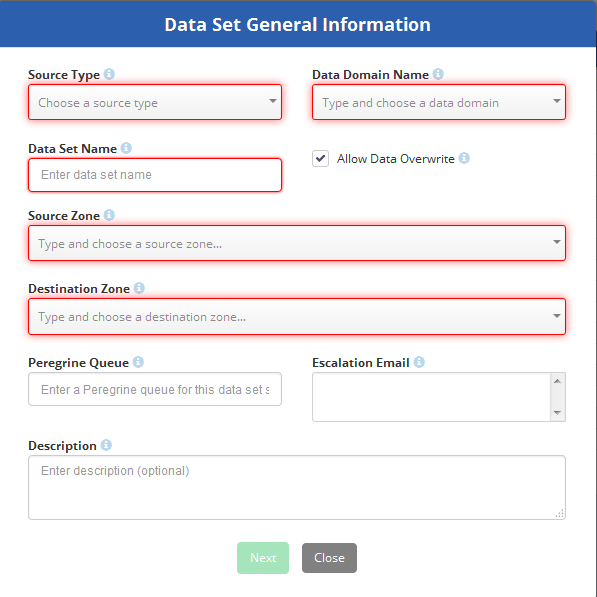
After clicking Yes, it will display a message if the File processing was successful or not.



A user can select the “Override” option if you wants to overwrite the existing data set over the new one with the same name.

# Create a Data Set

To create a data set, click on the  button located on top right of the Data set list page. A new pop up screen “Data Set General Information” appears

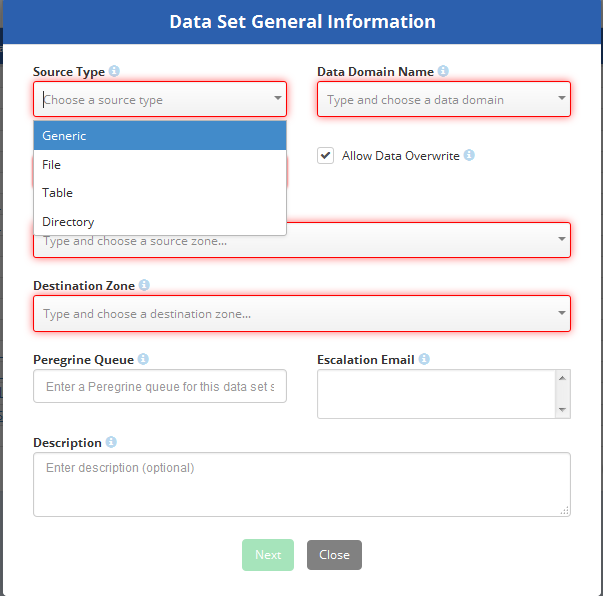


The field description and rules for registering a data set is the below table:

|  |  |  |
| --- | --- | --- |
| **Field Name** | **Description** | **Required** |
| **Source Type** | Data Set Source Name Generic File Table Directory | Y |
| **Data Domain Name** | Name of the Data Domain with objects to be added to policy | Y |
| **Data Set Name** | Available Data Sets under selected Data Domain | Y |
| **Allow Override** | Indicates whether the data can be writtened with the new data | Y |
| **Source Zone** | Resource used for the source of the data set | Y |
| **Destination Zone** | Resource to be used for the destination of the data set | Y |
| **Peregrine Queue** | Name of the Peregrine queue to send the incident Ticket to escalate resource connectivity issues | N |
| **Escalation Email** | Support email for issues related to this data set | N |
| **Description** | A brief description of the data set. | N |

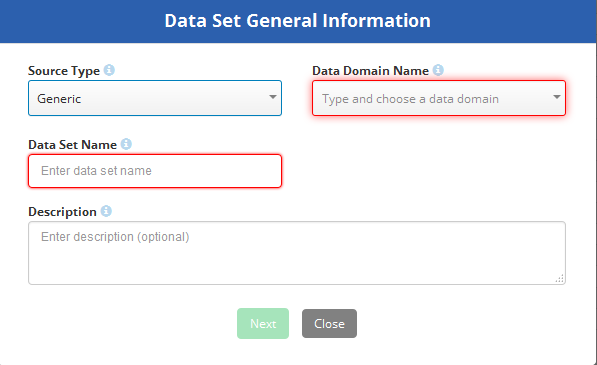
The source type can be of the following 4 types:

1. Generic
2. File
3. Table
4. Directory



# Source type: Generic

If the source type selected is Generic then the following screen appears

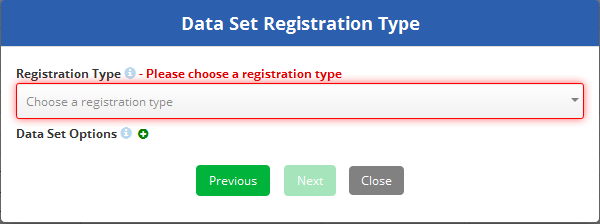


The field description and rules for registering a data set is the below table:

|  |  |  |
| --- | --- | --- |
| **Field Name** | **Description** | **Required** |
| **Source Type** | Data Set Source Name: Generic | Y |
| **Data Domain Name** | Name of the Data Domain with objects to be added to policy | Y |
| **Data Set Name** | Available Data Sets under selected Data Domain | Y |
| **Description** | A brief description of the data set. | N |

Once at least the Data Domain and Data set is filled in, click on the  button

Next select the Registration Types

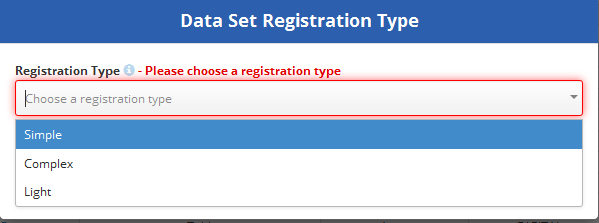


Registration types are a drop down menu and it can have values of:

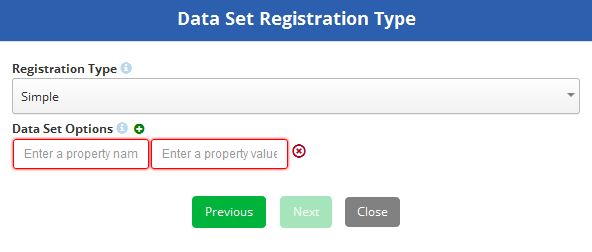
* Simple
* Complex
* Light

At least one value should be selected.

# Source type: Generic and Registration type simple



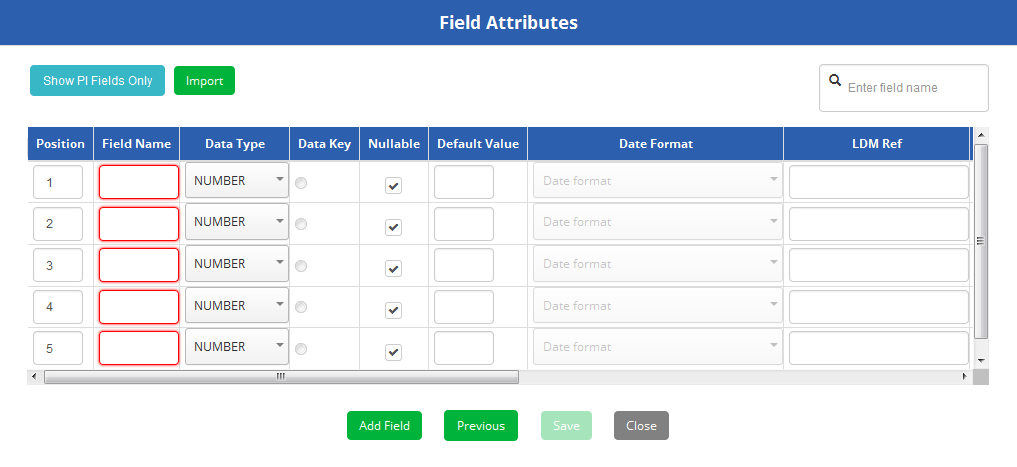
The data set options can be enabled by clicking on the  sign next to it. Below text area appears to fill the property name and property value for the data set.

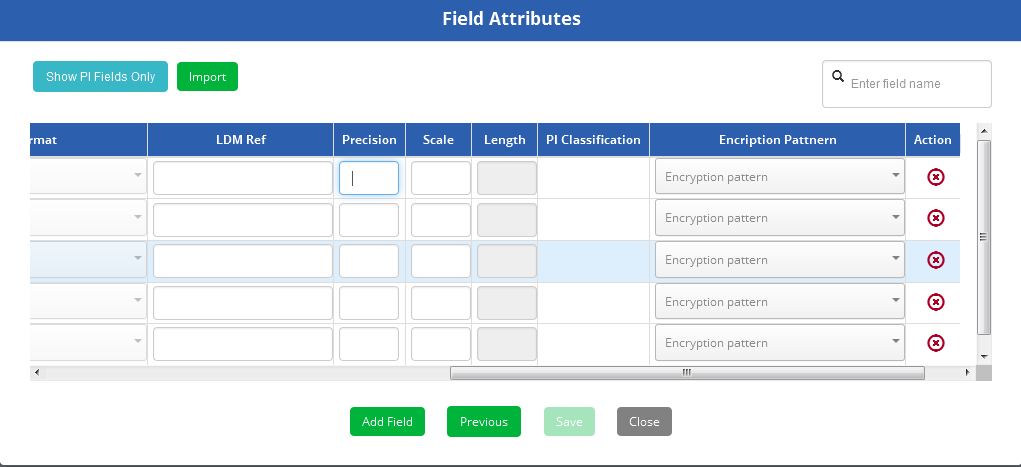


If you wish to cancel entering the Property Name and Property value , click on the  sign next to it.

Click on the  button to proceed further.

Enter the Field Names and select their corresponding data types. A user can also import





Enter the field name, data types, Default value, LDM Reference, precision (number only), scale (number only).

The following data types are supported:

* Number
* String
* Char
* Boolean
* Date
* Timestamp
* Binary
* Struct
* Array
* Map
* Others

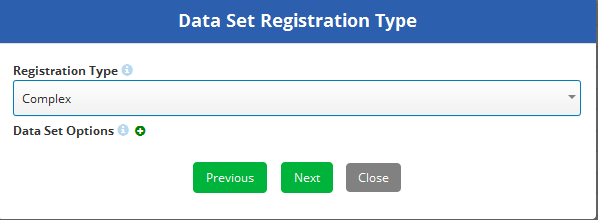
A new field can be added by click on the  Button

Fields entered or empty fields can be deleted by clicking on the  icon

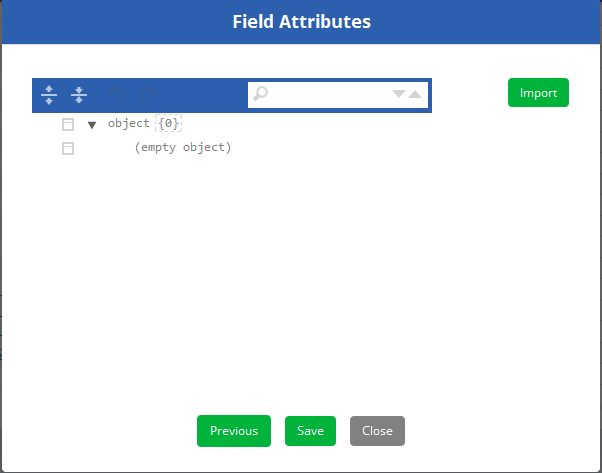
If a user wish to view the PI classified field, it can be enabled by clicking on the  button.

Once the entered field’s values look good, you can click on  button to create a data set.

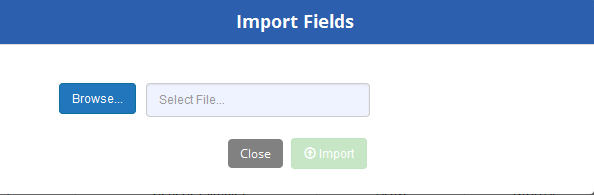
# Source type: Generic and Registration type simple



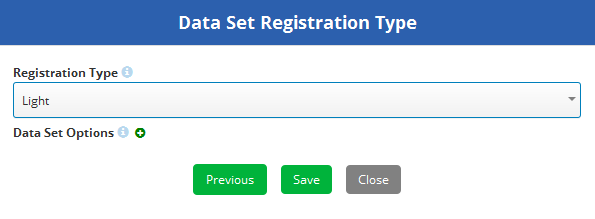
Select the Registration type as Complex and click on Next



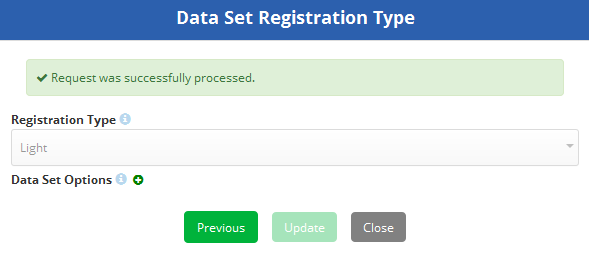
A user can click on the  button to import a field attributes



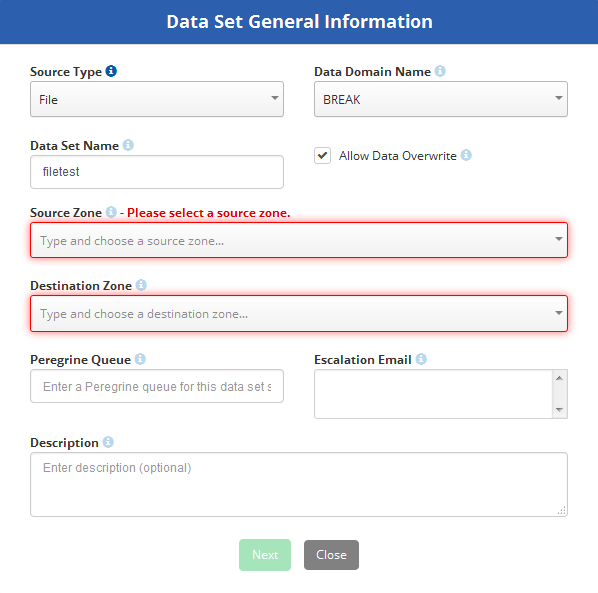
# Source type: Generic and Registration type light



Select light as the Registration type and click on Save. A message confirming request was successfully processed appears.



# Source type: File

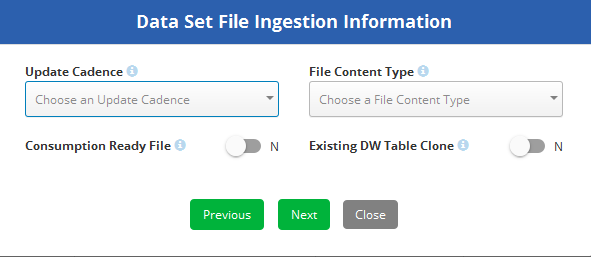


Select the source type as file and enter the Data Domain Name and enter a Data set name. If you wish to overwrite an existing data set, click on the “Allow Data Overwrite” option.

Select the source zone. Currently there can be 2 source zones

* Dropbox
* Hadoopconformed

Select the destination Zone. Enter the peregrine queue and Escalation Email and Description. This is optional. And click on next



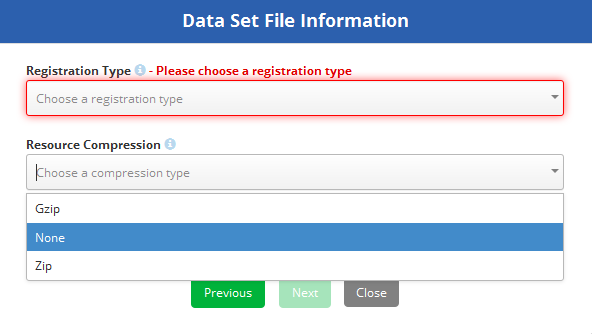
Select the Update Cadence value and File Content Type value and click next button(this is not mandatory). The following screen appears to enter the Registration Type and Resource Compressions.

The Regististration types can be of type

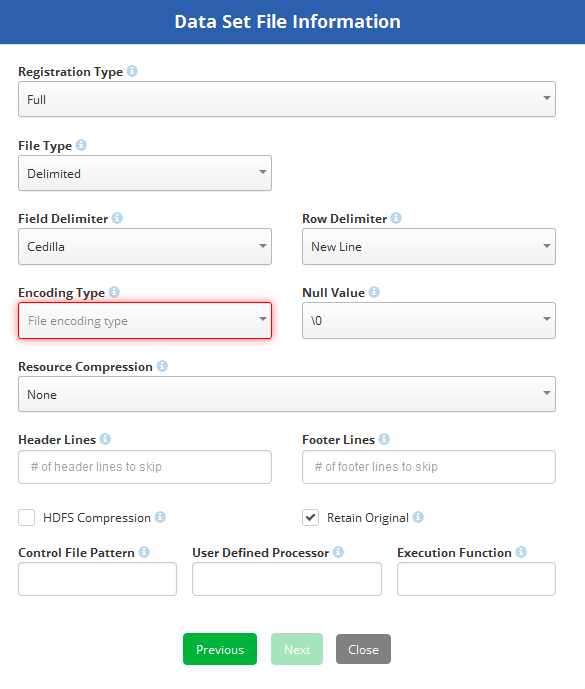
* Full
* Light

Resource compression can be of the type

* Zip
* Gzip
* None



# Registration type: Full

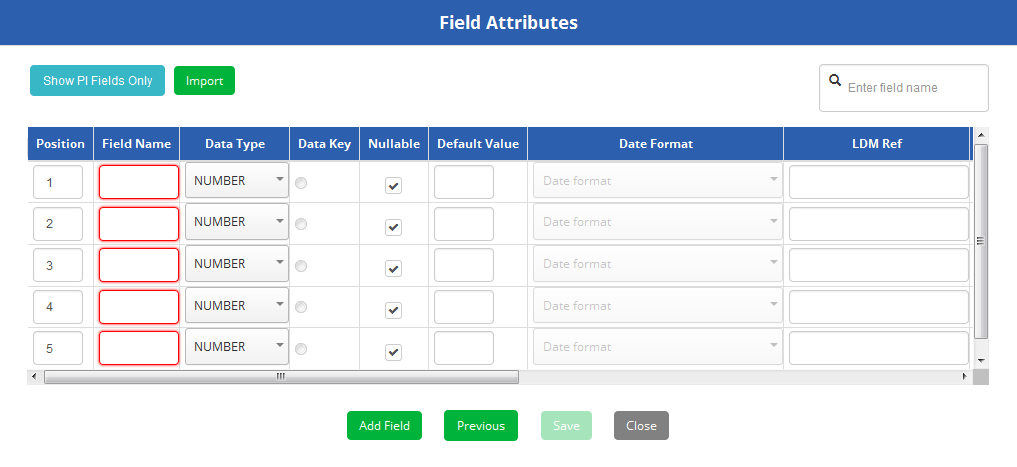


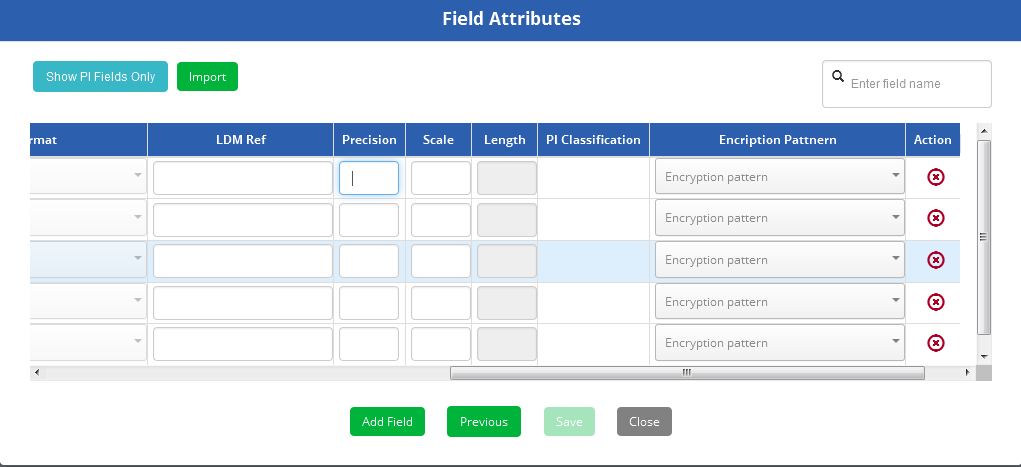
The field description and rules for registering a data set is the below table:

|  |  |
| --- | --- |
| **Field Name** | **Description** |
| **Registration Type** | Type of registration indicating if there are data fields to be registered within the data set. If the selected value is Fill |
| **File Type** | Type of file indicating the if the file is delimited or 2 dimensional |
| **Field Delimiter** | Type of delimiter that separates fields in a file. Values in drop down are Caret Cedilla Coma Pipe SOH Space Tab Tilde |
| **Row delimiter** | Type of delimiter that seperates rows in a file. Value in drop down New Line |
| **Encoding Type** | Encoding method used when the file is sent to the destination zone. Values in the drop down menu are: ASCII ORC UTF8 |
| **Null Value** | Type of values for the null fields in a file. Value in Drop down \0 BLANK(" ") |
| **Resource Compression** | Indicates if the file can be compressed. Values in dropdown are Gzip None Zip |
| **Header Line** | Number of header lines in the file |
| **Footer Line** | Number of footer lines in the file |
| **HDFS Compression** | Indicates if the file can be compressed in HDFC |
| **Retain Original** | Retain an original copy of the file in the archival zone. |
| **Control File Pattern** | Pattern for the control file to be used during ingestion |
| **User Defined Processor** | Script to be invoked for the ingestion control file preprocessing |
| **Execution Function** | Execution function to be used during ingestion |

Click on the  button to proceed once the information is entered.

Enter the Field Names and select their corresponding data types. A user can also import





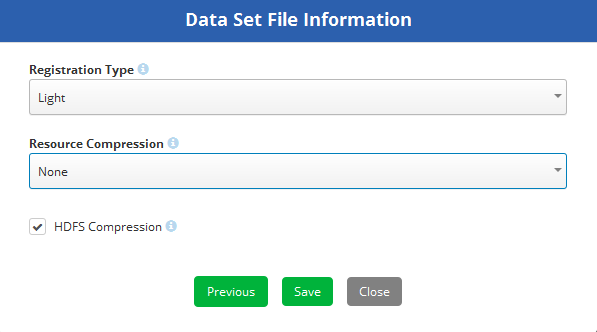
Enter the field name, data types, Default value, LDM Reference, precision (number only), scale (number only).

The following data types are supported:

* Number
* String
* Char
* Boolean
* Date
* Timestamp
* Binary
* Struct
* Array
* Map
* Others

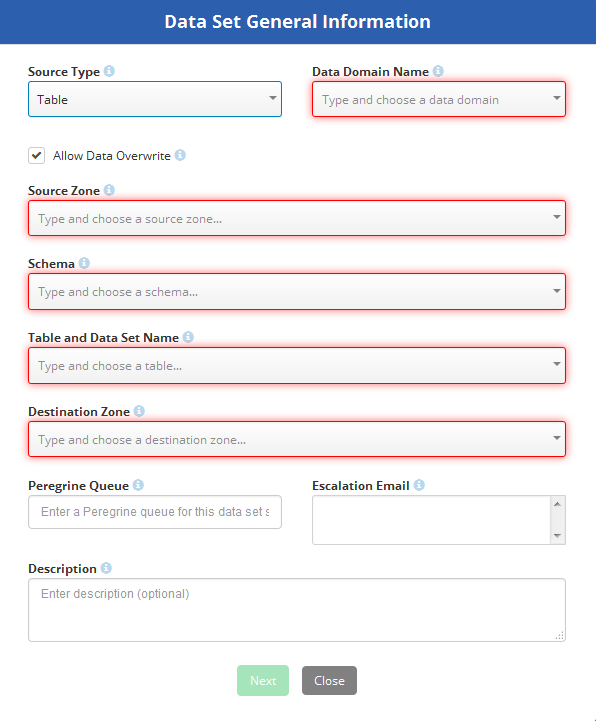
# Registration type: Light

Select the Registration type as light and resource compression as Zip, Gzip or Null and click on  the button



If the file needs to be compressed in HDFC, activate the “HDFC Compression” tick

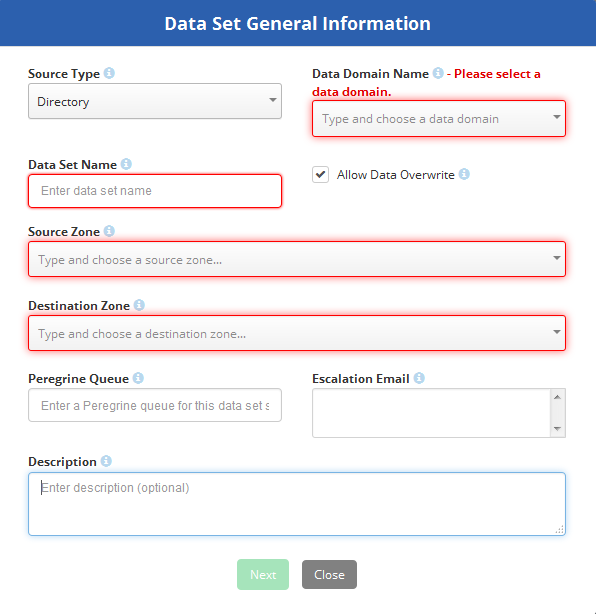
# Source type: Table



The field description is following:

|  |  |
| --- | --- |
| **Field Name** | **Support email for issues related to this data set** |
| **Source Type** | Source of the data set |
| **Data Domain Name** | Name of the data domain with objects to be added to the policy |
| **Allow Data Overwrite** | Indicates whether the data can be whitened with the new data |
| **Source Zone** | Resource used for the source of the data set |
| **Schema** | Database structure of the data set table |
| **Table and Data Set Name** | Name of the data set table |
| **Destination Zone** | A brief description of the data set. |
| **Peregrine queue** | Name of the Peregrine queue to send the incident Ticket to escalate resource connectivity issues |
| **Escalation Email** | Support email for issues related to this data set |
| **Description** | A brief description of the data set. |

# Source type: Directory



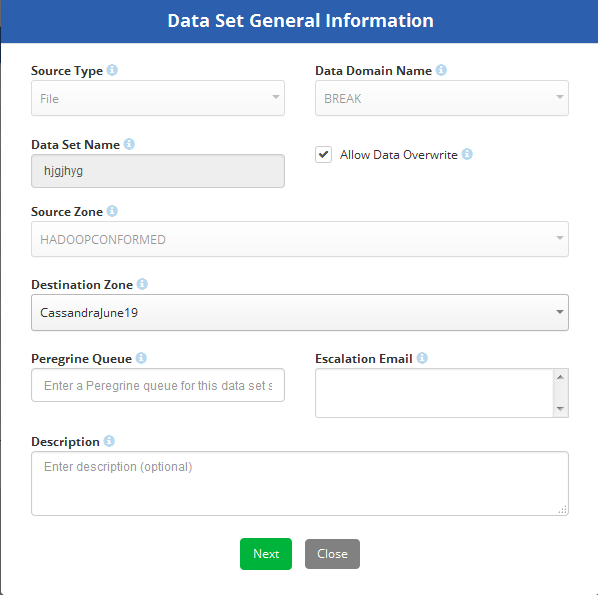
The field description is following:

|  |  |
| --- | --- |
| **Field Name** | **Support email for issues related to this data set** |
| **Source Type** | Source of the data set |
| **Data Domain Name** | Name of the data domain with objects to be added to the policy |
| **Allow Data Overwrite** | Indicates whether the data can be writtened with the new data |
| **Data Set Name** | Available data sets under selected data domain |
| **Source Zone** | Resource used for the source of the data set |
| **Destination Zone** | A brief description of the data set. |
| **Peregrine queue** | Name of the Peregrine queue to send the incident Ticket to escalate resource connectivity issues |
| **Escalation Email** | Support email for issues related to this data set |
| **Description** | A brief description of the data set. |

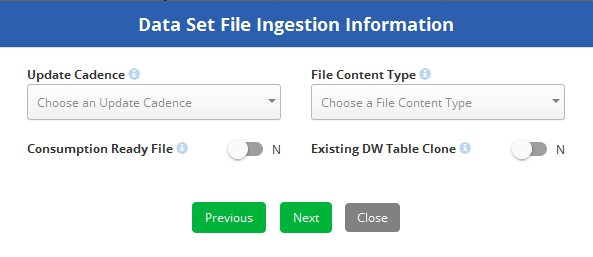
# Edit a Data Set

To edit a data set, click on the Data Set Name on the Data set list page. Depending on the type of source type the edit pop box would appear.

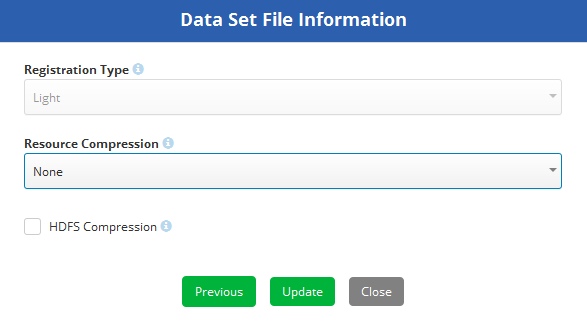
# Edit a Data Set l: file (Light)



Click next

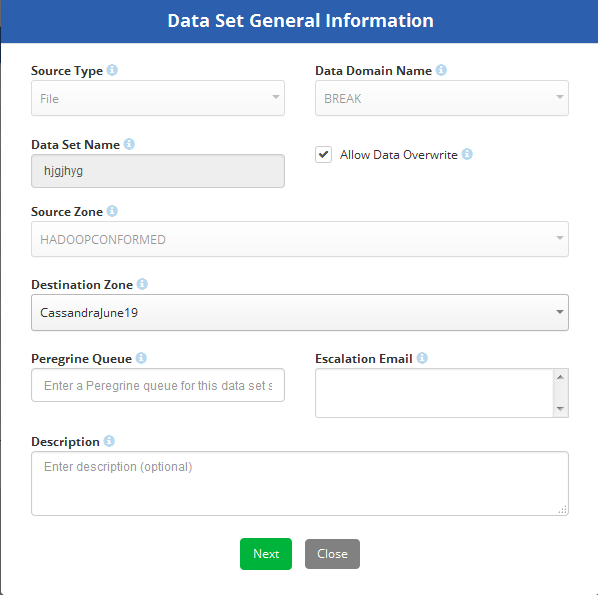


Make changes where applicable in this screen and click next

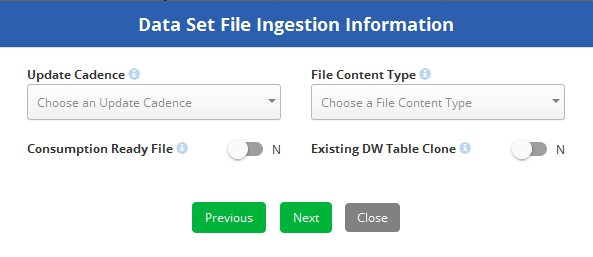


Make changes where applicable in this screen and click Update.

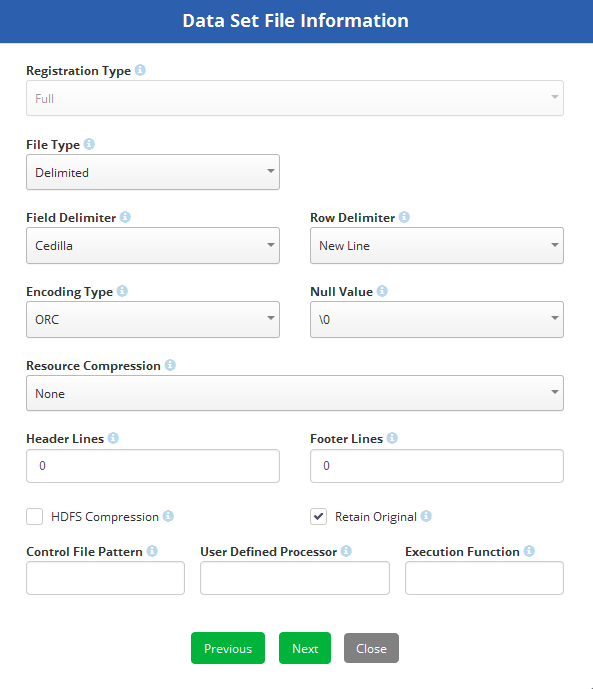
# Edit a Data Set l: file (Full)



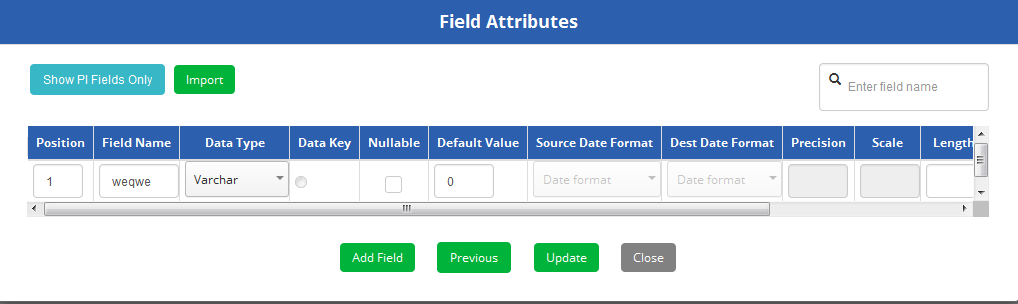
Click next



Make changes where applicable in this screen and click next

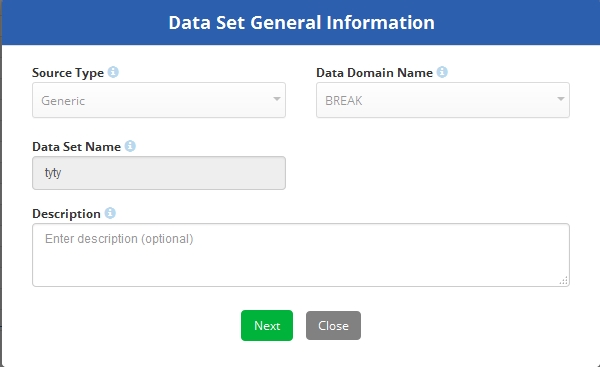
.

Make changes where applicable in this screen and click next

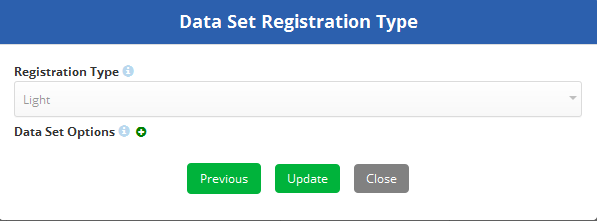


Make changes where applicable in this screen and click update.

# Edit a Data Set l: Generic (Light)

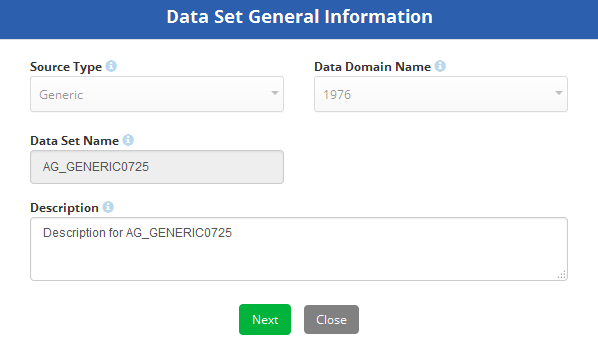


Make changes where applicable in this screen and click Next

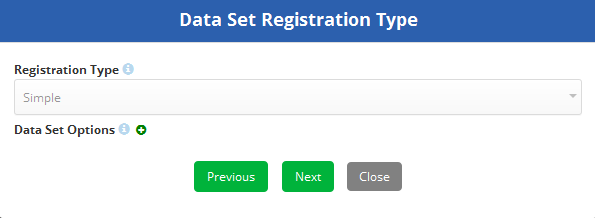


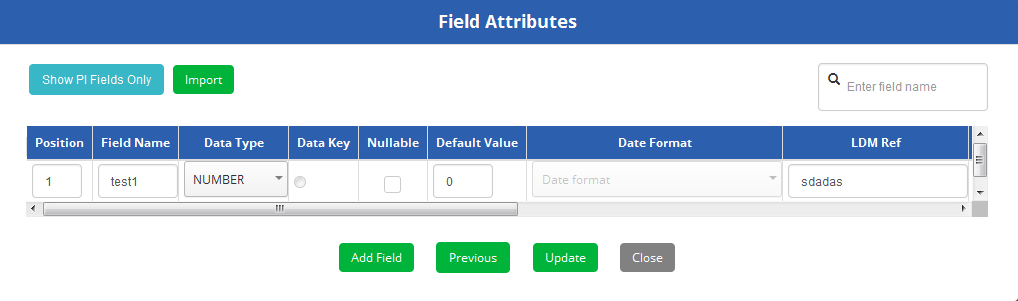
Make changes where applicable in this screen and click update.

# Edit a Data Set l: Generic (Simple)



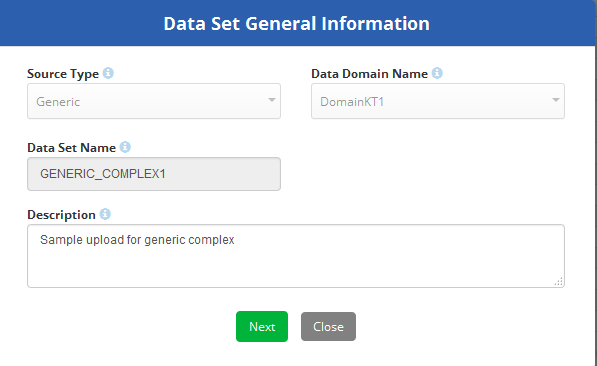
Make changes where applicable in this screen and click Next

  
Make changes where applicable in this screen and click Next

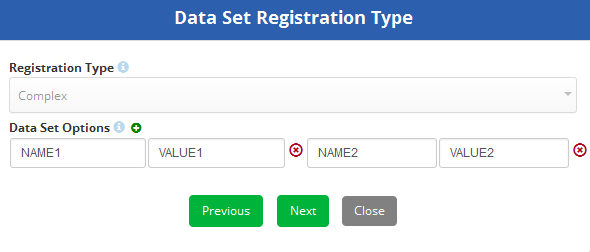


Make changes where applicable in this screen and click update.

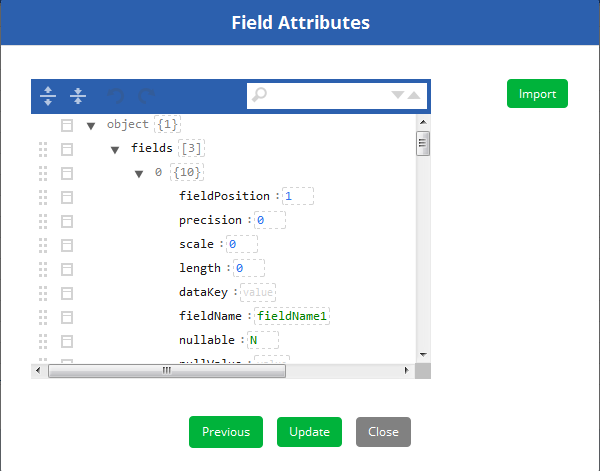
# Edit a Data Set l: Generic (Complex)



Make changes where applicable in this screen and click Next



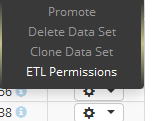
Make changes where applicable in this screen and click Next



Make changes where applicable in this screen and click update.

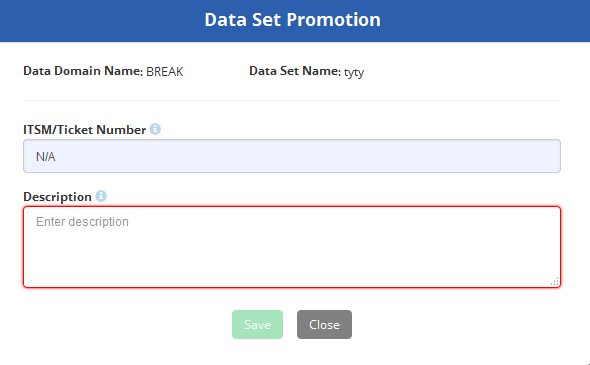
# Delete, promote, Clone or add ETL Permissions to a Data Set

In the data set list page, under the action column a user will see the  icon. For a data set, following actions are available under the action column



* Promote
* Delete Data Set
* Clone Data Set
* ETL Permissions

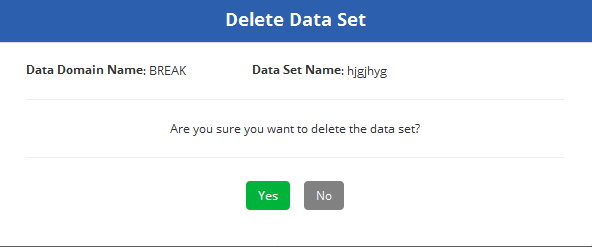
**Promote**: this promotes the data set from lower environment to higher environment but it cannot promote from higher to the lower environment. For example it can promote from design to integration and/or integration to preview but it cannot other way around. For some Data domains, a user may not see a promote feature available. This is because they need to a have correct LDAP roles to promote. Below table shows what LDAP roles are required to what type of promotions?



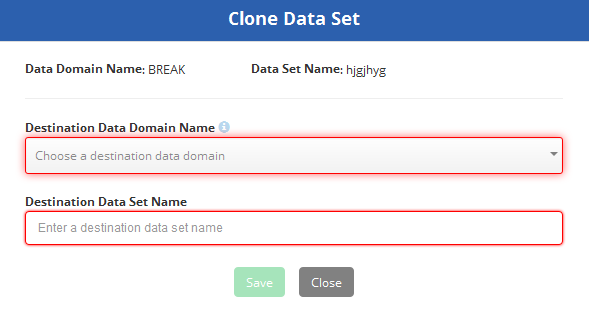
|  |  |  |
| --- | --- | --- |
| **Environment promotions** | **LDAP Role** | **LDAP Role Description** |
| From Design to Integration | UDS\_DEV\_META\_PROMOTER | Provide users ability to promote metadata for the following from Design to Preview: - data domain, datasets, encryption policies, data flows, data processors |
| From Integration to Preview | UDS\_QA\_META\_PROMOTER. | Provide users ability to promote metadata for the following from Integration to Preview: - data domain, datasets, encryption policies, data flows, data processors |

**Delete:** A data set can be deleted as long as it is not used by any processors

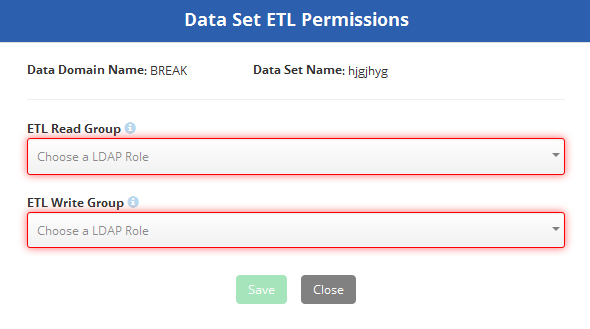
A dialogue box appears asking approval to proceed with deleting the data set as shown in the below figure



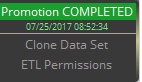
**Clone Data Set:** This option is to clone an existing data set. A destination data domain name and destination data set name should be selected where this data set clone needs to be put into.



**ETL Permissions:**



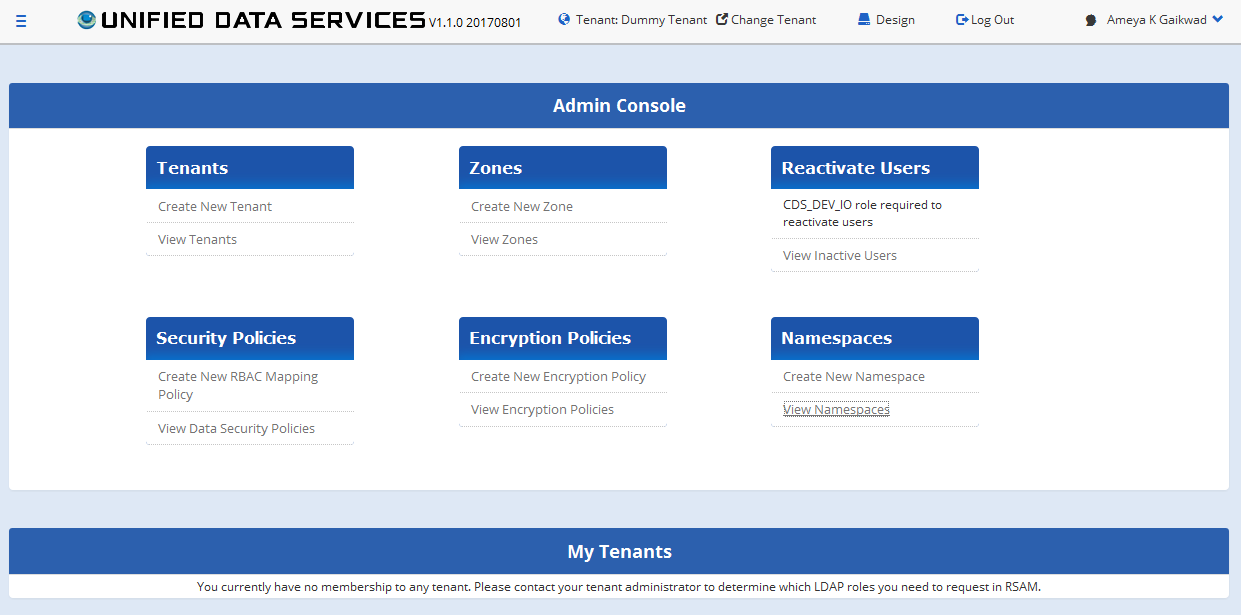
Some data sets will have a  icon under the action column rather than the conventional  icon. The green icon indicates that this particular data set has already been promoted.



# Namespaces

# Namespace List Page

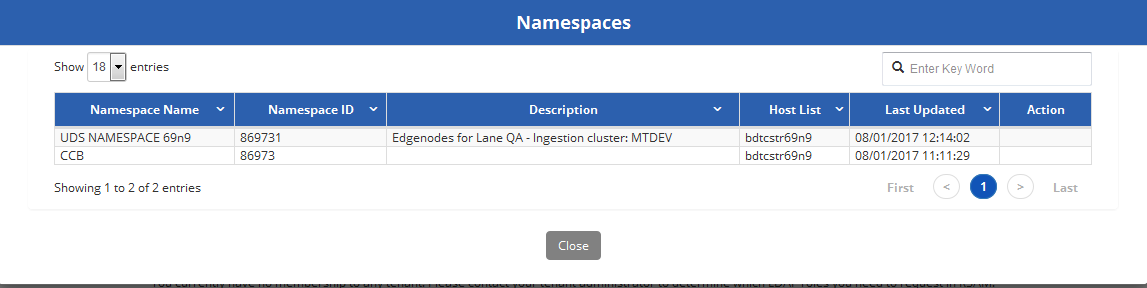
The namespace list page can be access via the main menu.



Under the main menu, a user can

1. Create a new name space
2. View Namespace
   1. Edit Namespace
   2. Delete Namespace

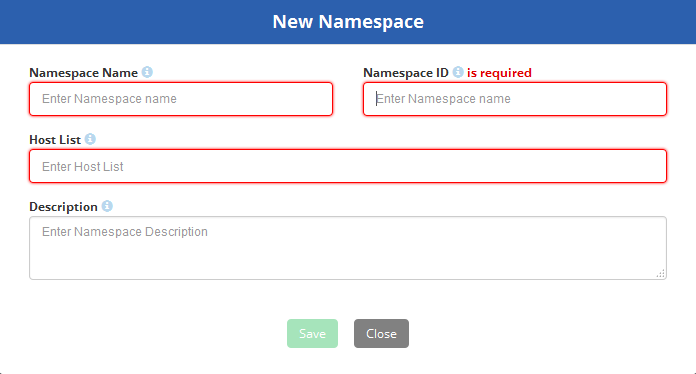
To view the namespace list page, click on ‘View Namespaces’. A pop up will appear displaying the following list page as show in the below image.



|  |  |
| --- | --- |
| **Field Name** | **Field Description** |
| Namespace Name | Name of the Namespace |
| Namespace ID | Unique ID used to identify a namespace. |
| Description | Description of the namespace |
| Host List | Server Name |
| Last Updated | Last updated timestamp |
| Action | Action to delete a namespace |

# Create a new Name space

To create a new name space, click on the ‘Create New Namespace’. A pop up will appear as shown in the below figure



|  |  |  |
| --- | --- | --- |
| **Field Name** | **Field Description** | **Required?** |
| Namespace Name | Name of the Namespace. Accepts timestamp format. | Y |
| Namespace ID | Accepts timestamp format. | Y |
| Host List |  | Y |
| Description | Description of the namespace | N |

# Edit a Namespace

To Edit a namespace which has been created in the past, go on the Namespace list page and click on the Namespace name you wish to edit. Following is the list of fields which shows which can be and which cannot be edited as part of Namespace.

|  |  |
| --- | --- |
| **Field Name** | **Editable?** |
| Namespace Name | Y |
| Namespace ID | N |
| Host List | Y |
| Description | Y |

# Delete a Namespace

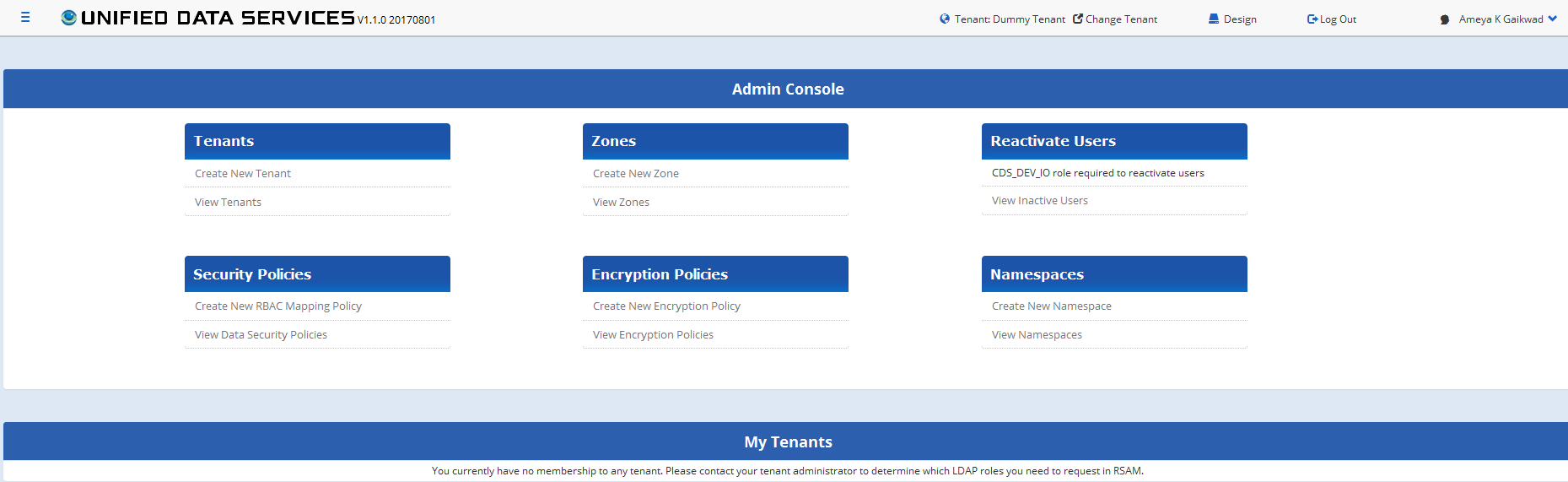
To delete a namespace, go on the namespace list page and under the action column, select the name space you wish to delete

**Note**: If a namespace is referenced by a tenant, it cannot be deleted.

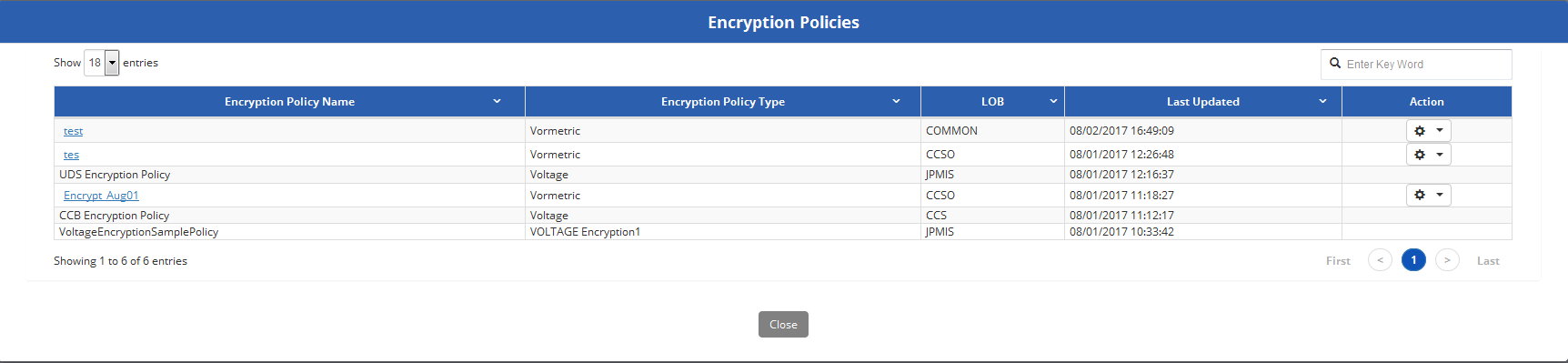
# Encryption Policy

# View Encryption Policy

To view encryption policies click on the ‘View Encryption Policies’ card on the UDS home page.



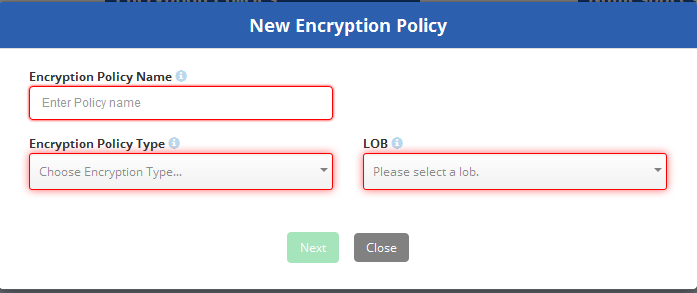
A pop up window will appear displaying the encryption policies. By default the latest 18 Encryption policies will be displayed



# Create an Encryption Policy

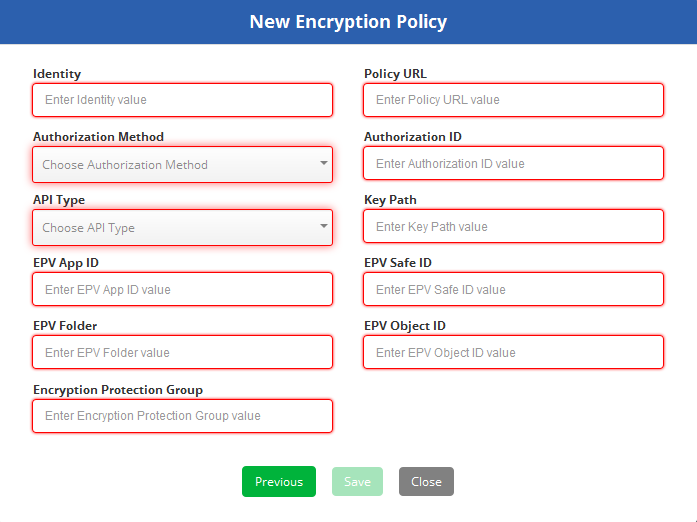
To create a new Encryption policy click on the ‘Create New Encryption Policy’ button on the UDS home page under the ‘Encryption’ card.

A pop up window appears to create a new encryption policy as shown in the below image.



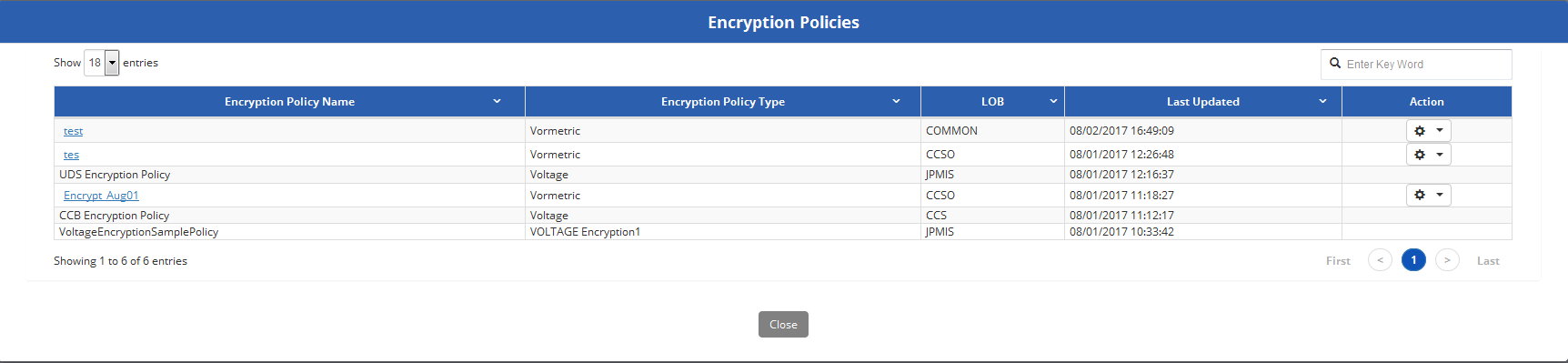
|  |  |  |
| --- | --- | --- |
| **Field Name** | **Description** | **Required?** |
| Encryption Policy Name | Unique name of the encryption policy. Is this unique globally or across tenants? | Y |
| Encryption Policy Type | There are 2 types of Encryption policies: Voltage Vormetric | Y |
| LOB | Line of Business | Y |

Irrespective of the selection type of ‘Encryption Policy Type’, once all the information is added, click on the ‘Next’ button.

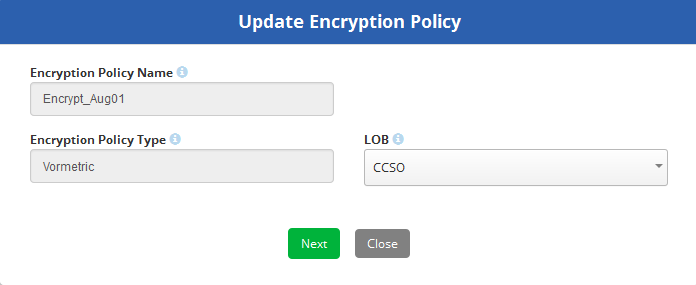


# Edit or Update an Encryption Policy

To edit an encryption policy, click on the policy name on the list page.

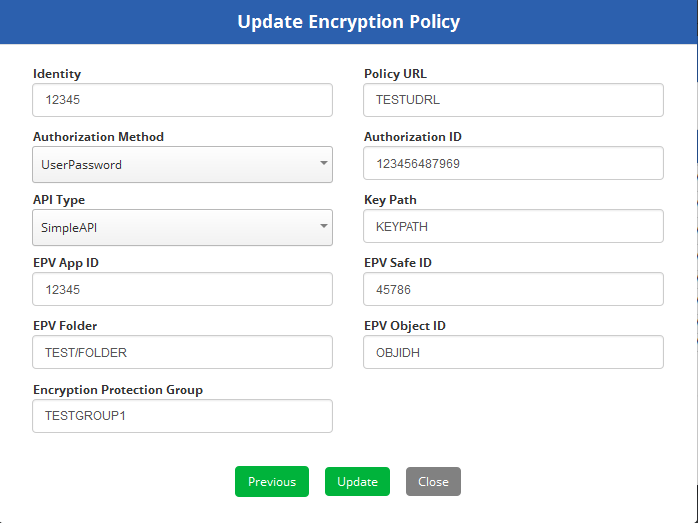


An ‘Update Encryption Policy’ pop up window will appear as shown in the below image.



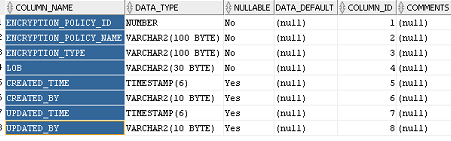
‘Encryption Policy Name’ & ‘Encryption Policy Type’ cannot be edited. Only ‘LOB’ can be edited.

Once the new LOB is selected, click on ‘Next’. A new pop up window will appear as shown in the below image.

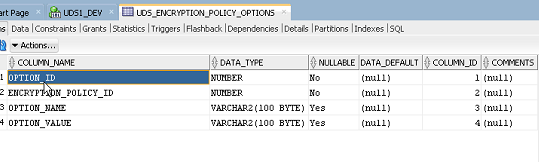


Update the fields and click ‘Update’ button

UDS\_ENCRYPTION\_POLICY

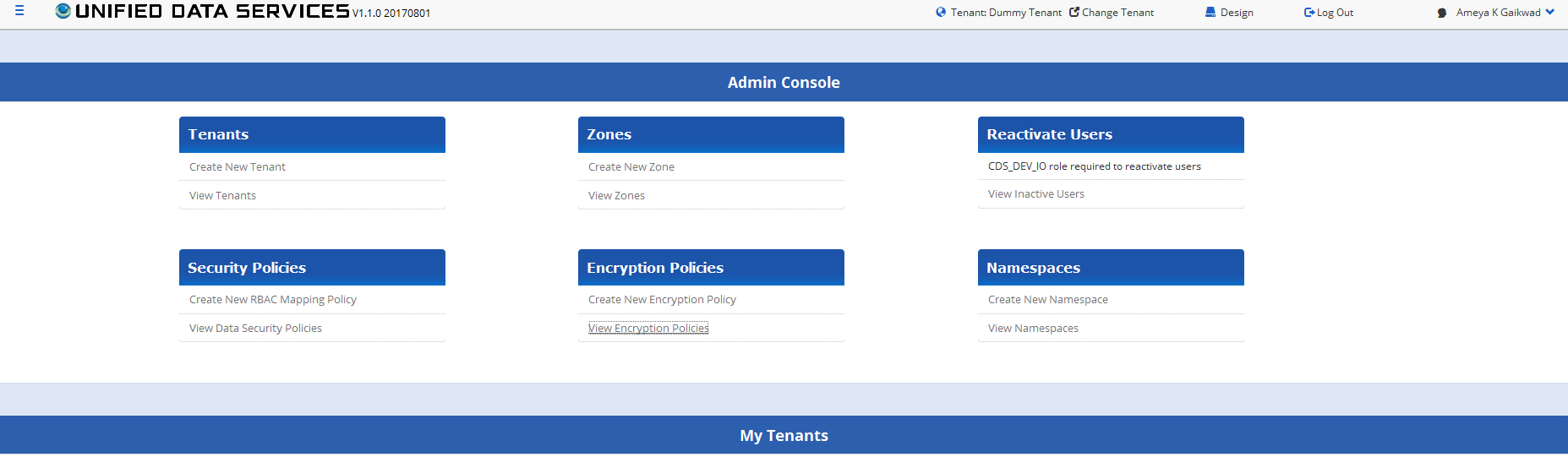


UDS\_ENCRYPTION\_POLICY\_OPTIONS

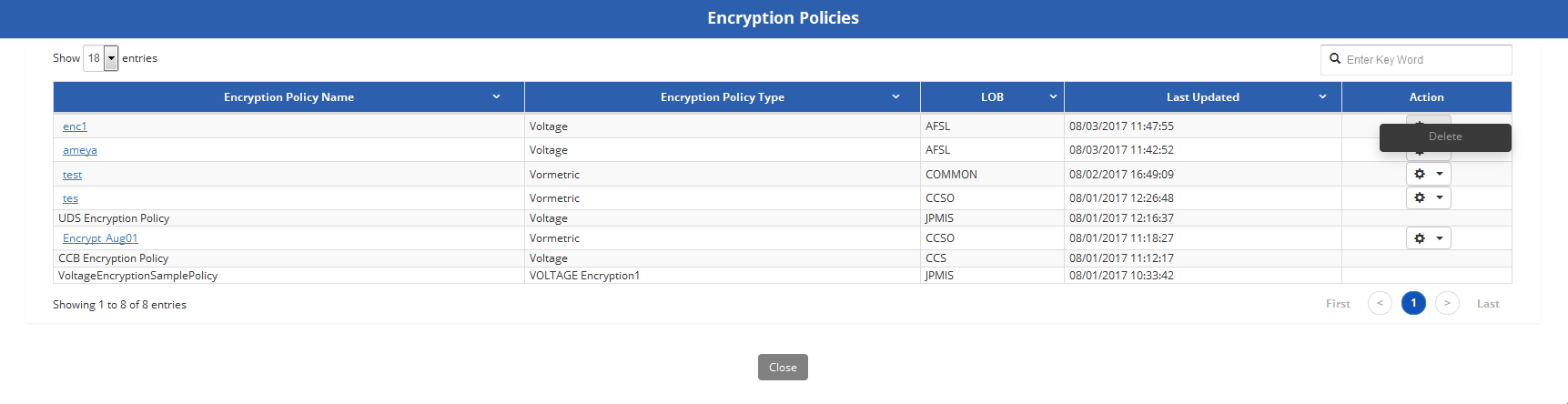


# Delete an Encryption Policy

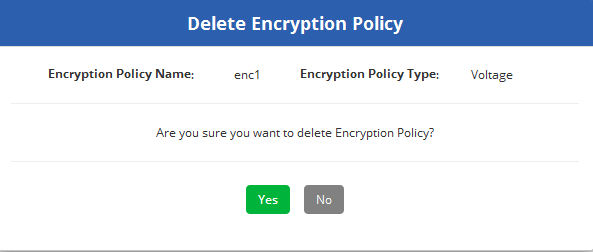
To delete an Encryption Policy, click on the ‘View Encryption Policies’ button on the UDS home page under the ‘Encryption’ card.



Select the Encryption policy you wish to delete on the list page and go to the ‘Action’ column and click on the  icon. A ‘Delete’ option will show up.



After the ‘Delete’ is selected a confirmation page will pop us asking the user if he wants to proceed with the deletion of the encryption policy



‘Yes’ will confirm the delete and ‘No’ will take the user back the Encryption policy list page.

**Note**: If an encryption policy is linked to a tenant, it cannot be deleted. So for some encryption policies, you may not see  icon next to it. This indicates that the particular encryption policy is linked to a tenant.

# Processors

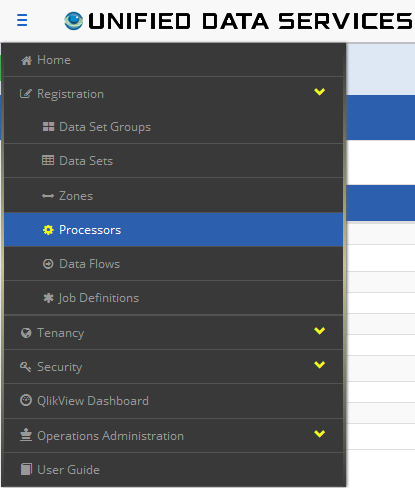
Processors are predefined processes to move the data from the source to target. There are 2 types of processors

* UDS: System created processors.
* UDP: A user created processor. A user can only create a UDP processor and not a UDS processor

# Processor list page

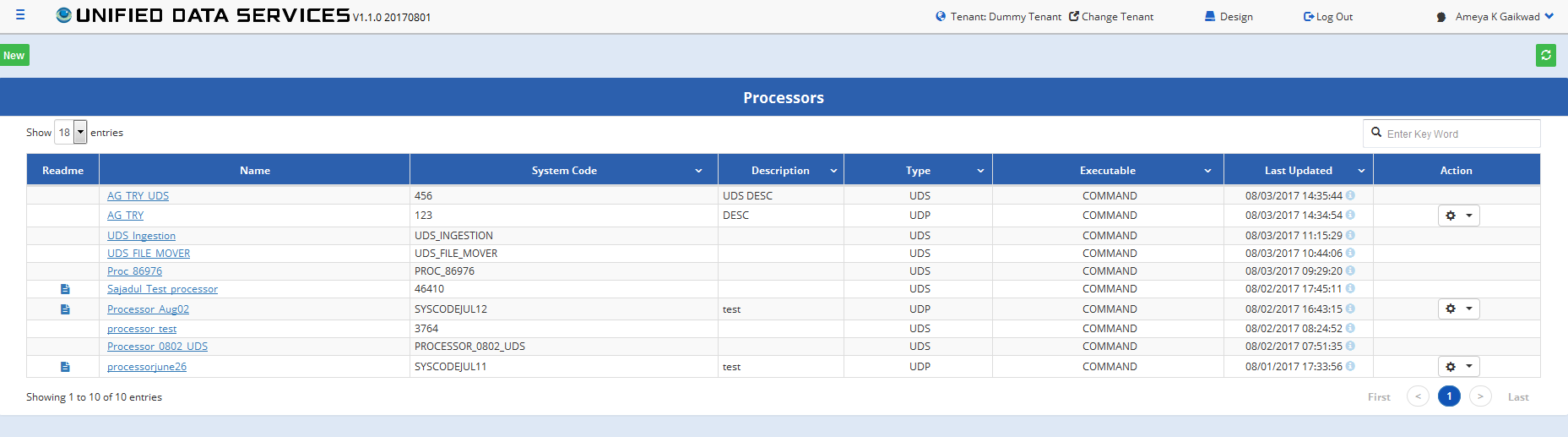
Processor list page shows all the predefined processors which have been created in the past. By default the processor list page shows latest 18 processors which have been created in the past sorted in descending order of their create date.

To Validate to the processor list page, click on the main menu on the UDS home page as shown in the below figure.



In the main menu on the top left side, you will see the link to the Processors under ‘Registration’.

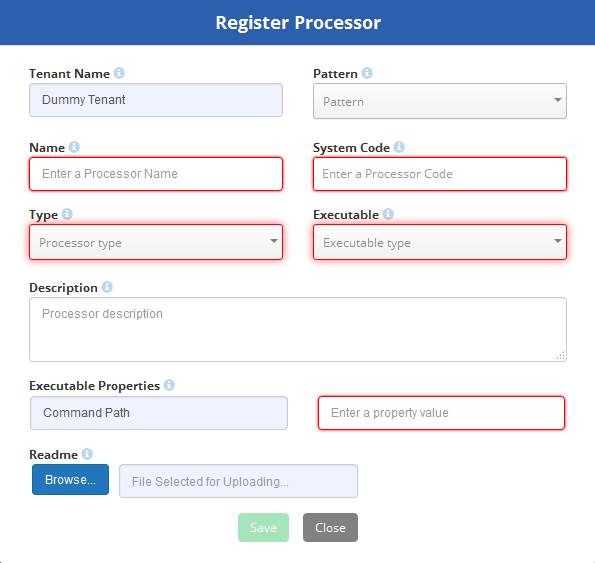
After selecting ‘Processors’ it will navigate to the processors list page as shown in the below image.



|  |  |
| --- | --- |
| **Field Name** | **Description** |
| Readme | A readme file to describe the processor. Only accepts .txt file formats. |
| Name | Unique processor name. Across tenant or Universally Unique?? |
| System Code | System code for the processor |
| Description | A brief description of the processor |
| Type | Available processor types. There are 2 types of processors UDS UDP |
| Executable | Available type of executable. For 1.1 it is only 'Command' |
| Last Updated | Last updated timestamp in the format of MM/DD/YYYY HH:MM:SS |
| Action | Action button dropdown to delete a processor. |

# Create a new Processor

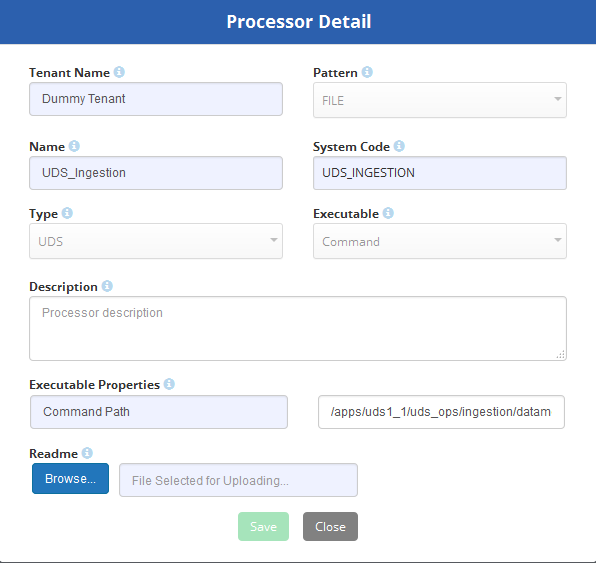
To create a new processor, click on the  button on the top left of the processor list page. A ‘Register Processor’ pop up page will appear as shown in the below image.



|  |  |
| --- | --- |
| Tenant Name | This name will be auto populated and is non editable. This will show the tenant name under which the user has logged in |
| Pattern | Following 3 types of patterns are supported  UDS  UDP  FILE |
| Name | Unique processor name. Across tenant or Universally Unique?? |
| System Code | System code for the processor |
| Type | Available processor types. There are 1 types of processors UDS |
| Executable | Available type of executable. For 1.1 it is only 'Command' |
| Description | A brief description of the processor |
| Executable properties | Command path to executable property |
| Readme | A readme file to describe the processor. Only accepts .txt file formats. A user can upload a readme file as part of Readme and this will be visible in the list page |

# Edit a Processor

To edit a processor, click on the ‘Name’ of the processor you want to edit in the processor list page. The following page will appear to edit the processor.



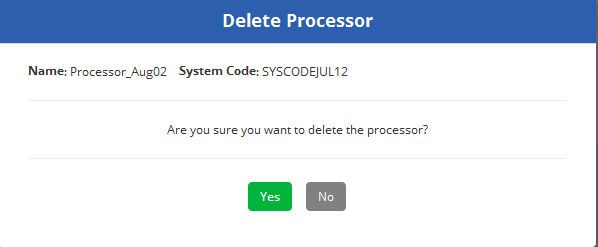
Only the following fields can be edited.

* Description
* Executable properties
* Readme file

# Delete a Processor

To delete a processor, navigate towards the Processor list page and click on the  icon under the action column. A drop down will appear and click on the  option.

After the ‘Delete’ is selected a confirmation page will pop us asking the user if he wants to proceed with the deletion of the processor.



‘Yes’ will confirm the delete and ‘No’ will take the user back the Encryption policy list page.

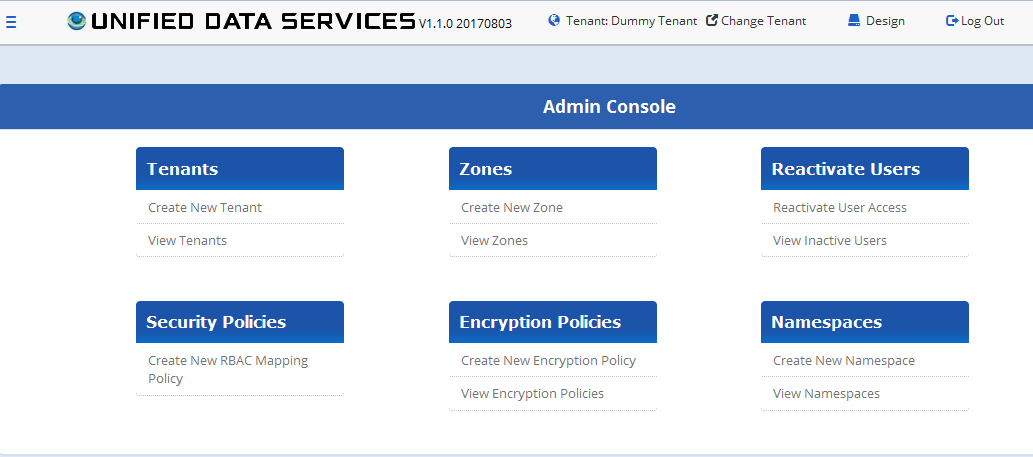
**Note**: UDS processors cannot be deleted. A UDP processor can be deleted. So UDS processors may not have  icon next to it. This indicates that the particular processor is UDS.

# Zones

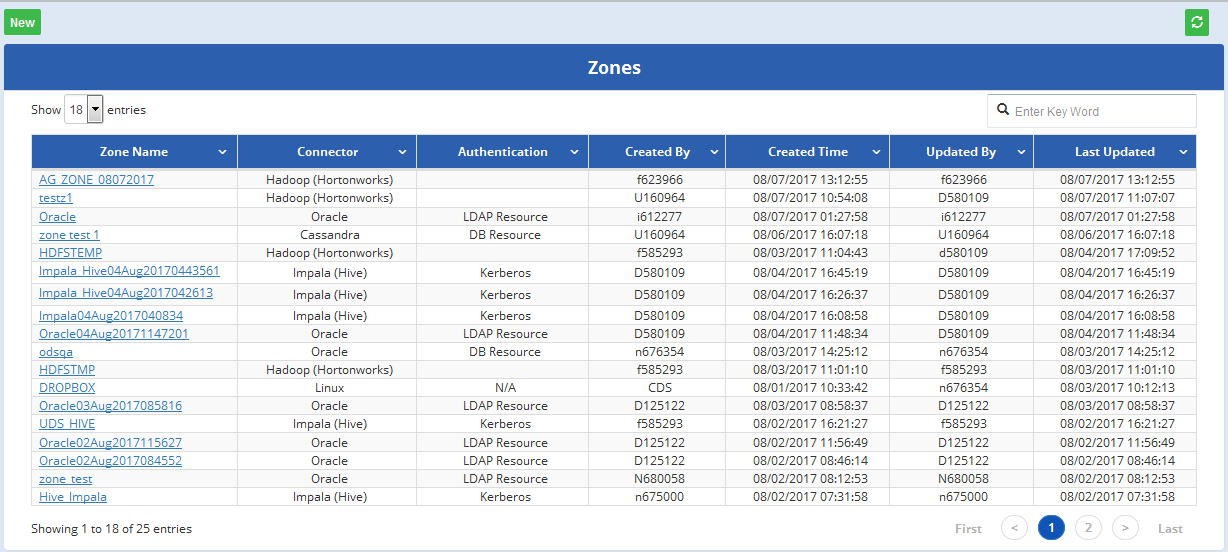
A Zone is defined by a system administrator and helps create a layer of abstraction over the underlying resource connectivity. Zone registration captures metadata around the physical container the data needs to be transmitted to or acquired from and the connectivity details that the UDS service will use to connect to it. It also helps the UDS team to switch out a particular container for another (more efficient) one transparently, without impacting the user.

# Zone list page

Zone list page displays all the zones which were created. In the UDS home page, click on the ‘View Zones’ under the ‘Zone’ card

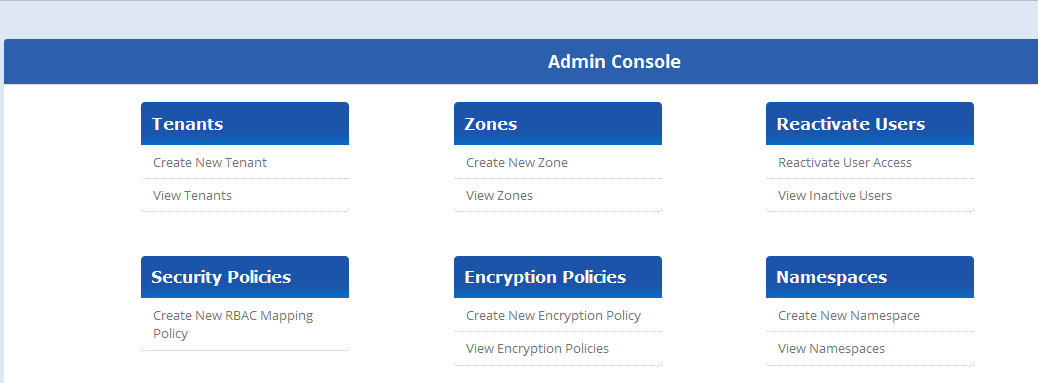


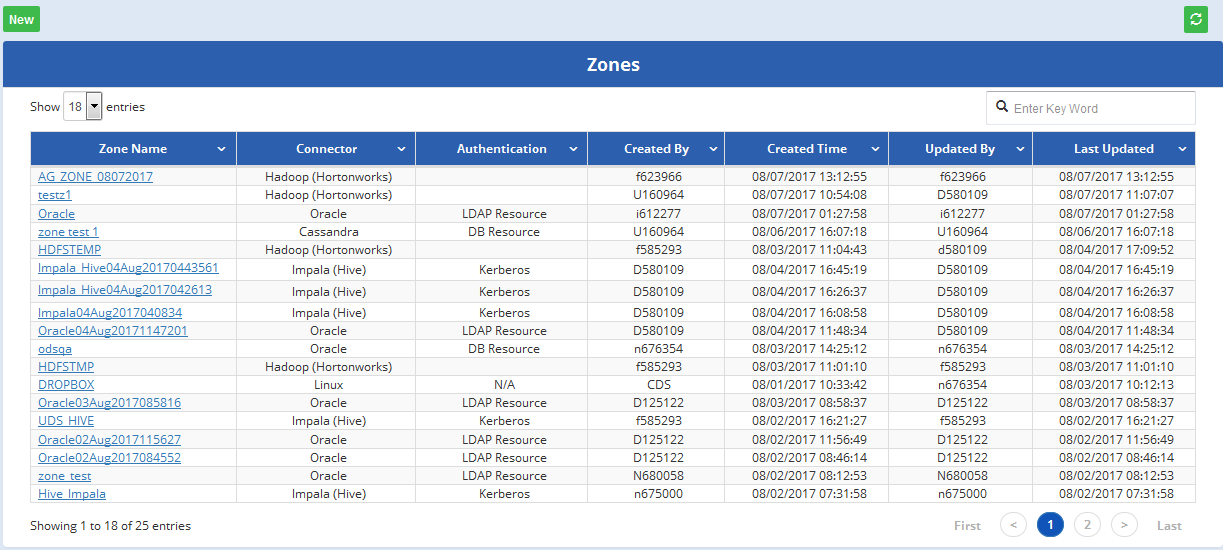
A new page will appear displaying the latest 18 zones which were created as shown in the below image.



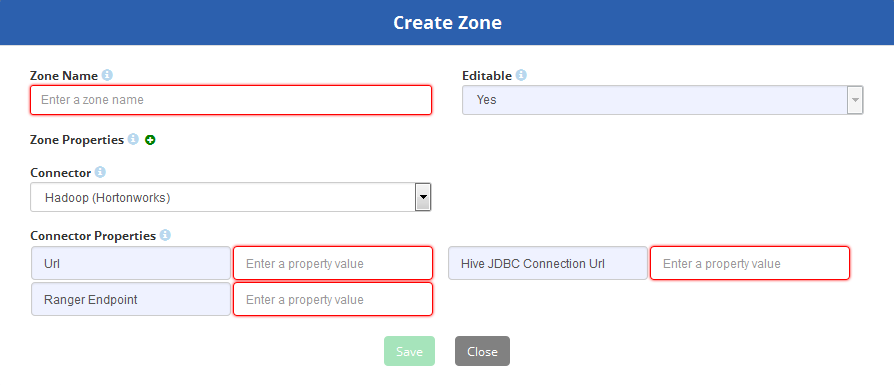
# Create a new Zone

To create a new zone, click on the ‘Create New Zone’ link under the ‘Zone’ card as shown in the below image.

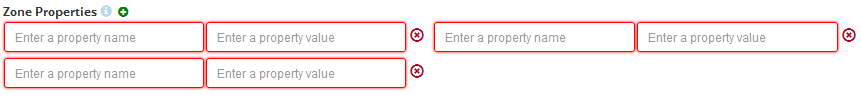
A user can also click on ‘New’ button on the ‘View zone’ list page to create a new zone as shown in the below image.



A new window appears to create a new zone as shown in the below image.



The text area marked in red is mandatory. A user can click on the sign next to the ‘Zone Properties’ to add the Property Name and Property Value. A user can add multiple sets of Property Name and Property Value as shown in the below image. Zone properties (Property Name and Property Value) is not a mandatory fields)

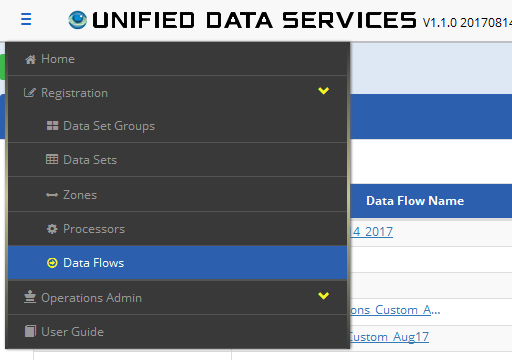


Once all the mandatory fields have, the ‘Save’ button will be enabled. Click on ‘Save’ to create the zone.

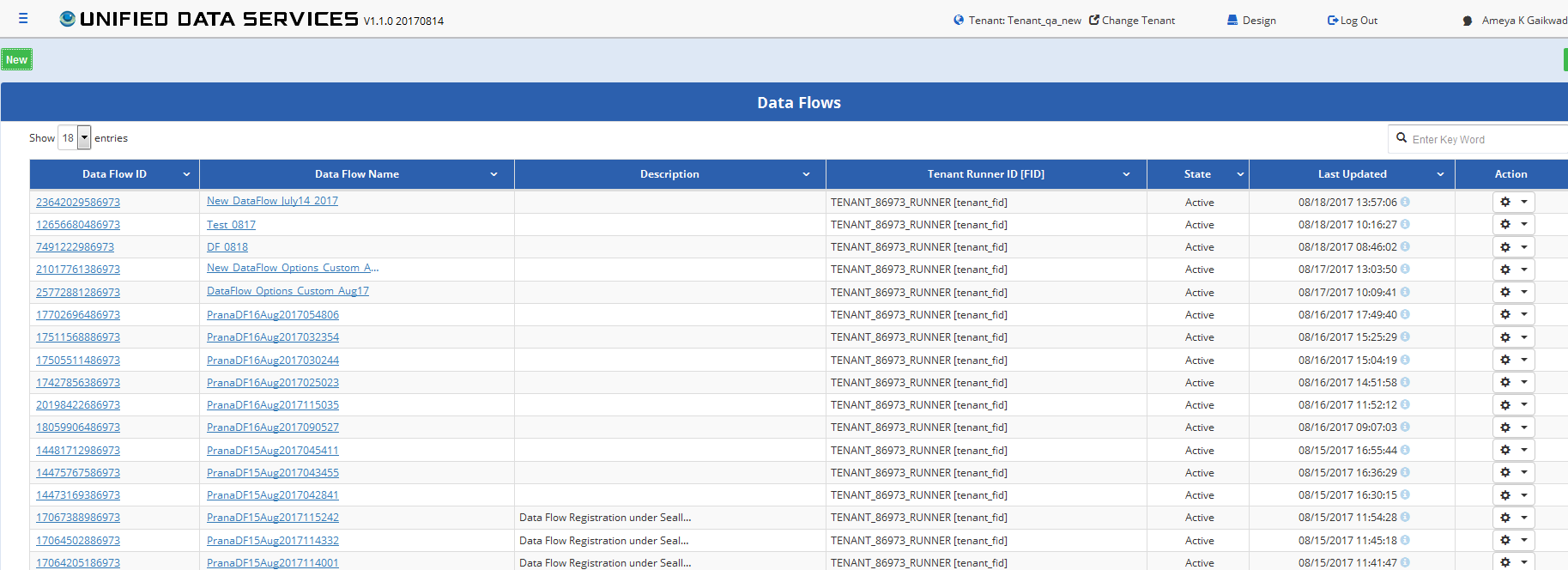
# Data Flows

# Data Flow List Page

Data flow is meta data about the execution process for the ingestor. To navigate to the data flow list page, a user needs to login under a specific tenant and click on the top left menu bar. “Data Flow” option will be visible under Registration.



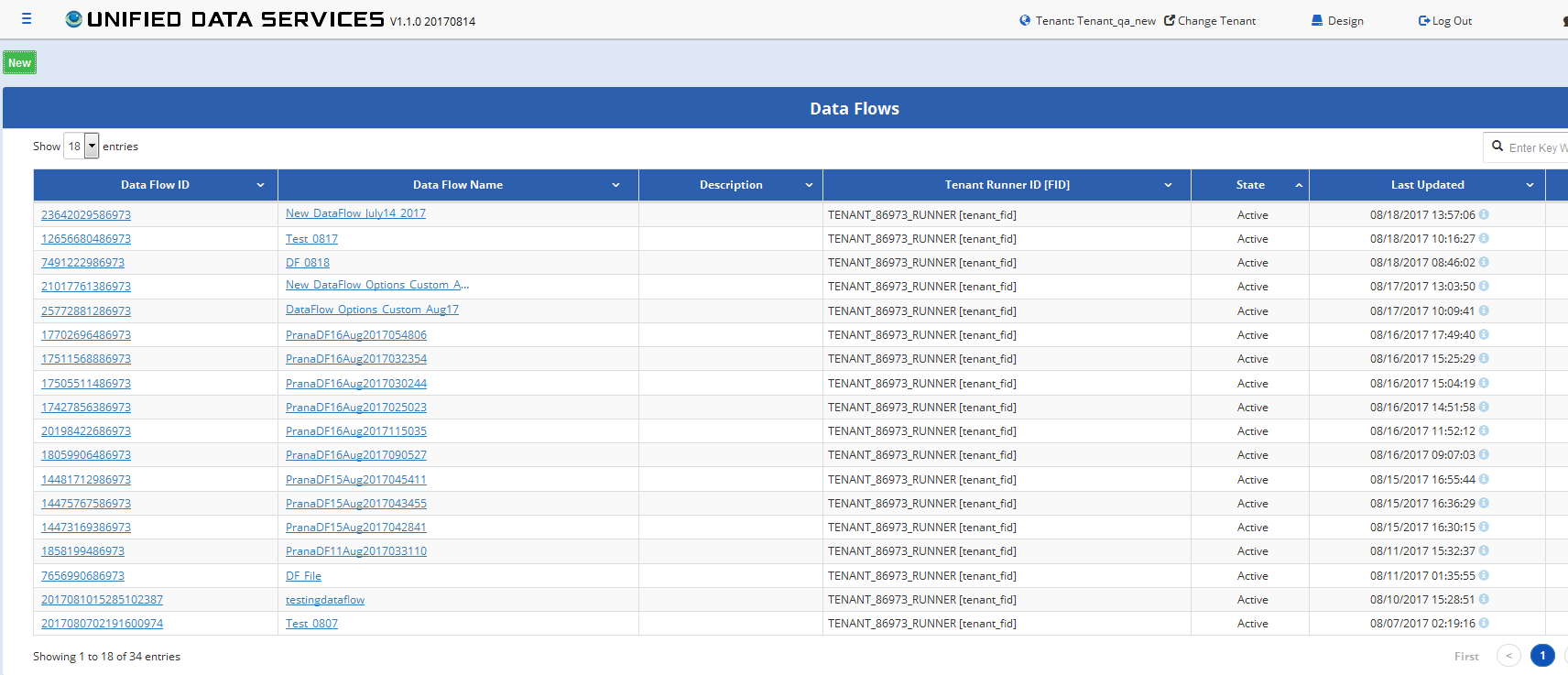
The data flow list page contains all the data flows which have been created in the past for that tenant. Below is the screenshot of the data flow list page.



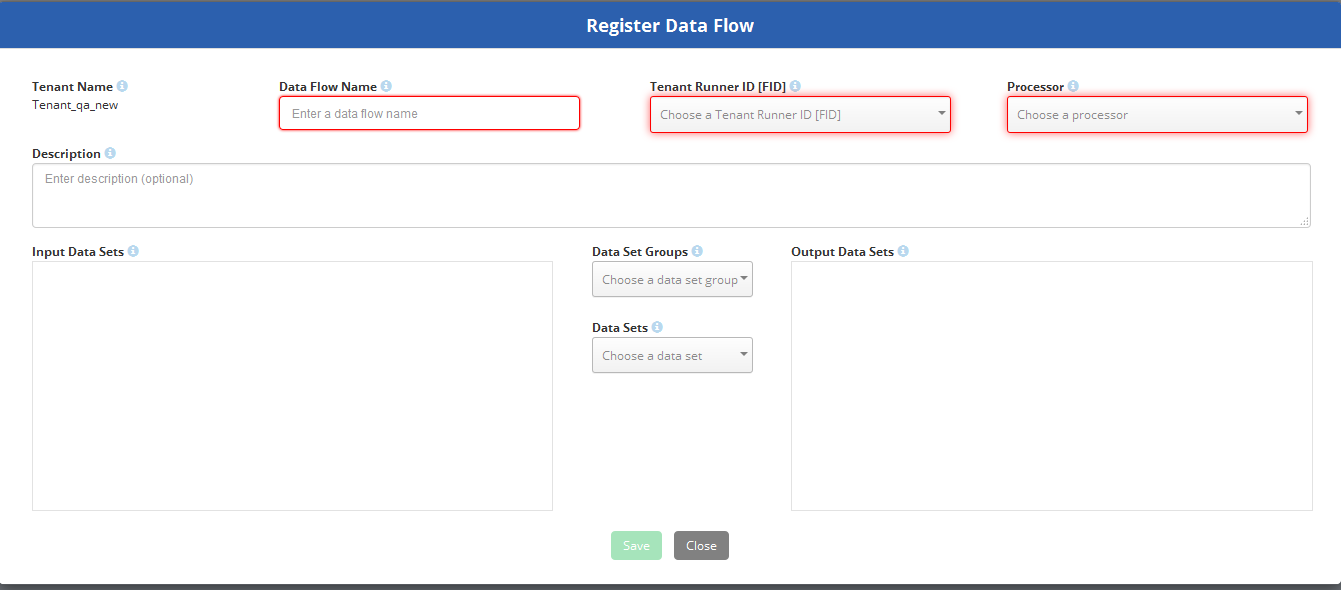
|  |  |
| --- | --- |
| **Field Names** | **Field Description** |
| Data Flow ID | A unique ID for the data flow |
| Data Flow Name | A unique name for the data flow |
| Description | Data Flow Description |
| Tenant Runner ID (FID) | Tenant Runner ID (FID) FunctionID |
| State |  |
| Last Updated | Last updated timestamp |
| Action | The actions available for a data flow is:  **Promote**: This promotes the data flow from a lower environment to a higher environment.  **Delete**: This option is used to delete a data Flow |

# Create a New Data Flow

To create a new data flow, click on the  button on the top left on the data flow list page.

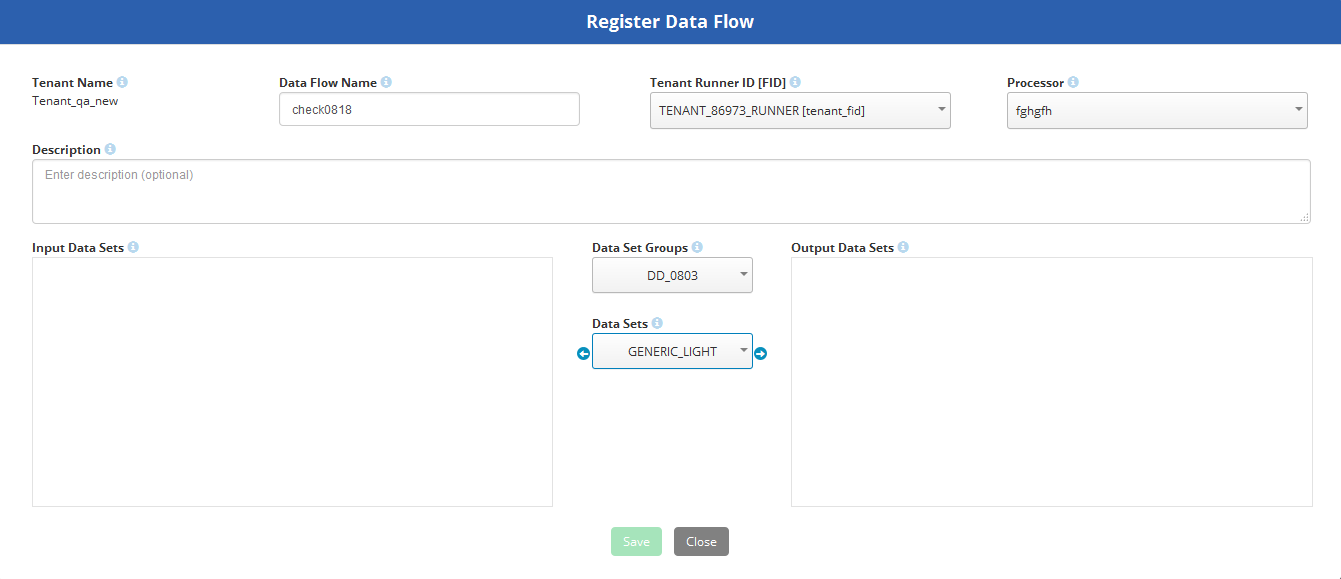


A new pop up appears to Register a new data flow appears as shown in the below image.



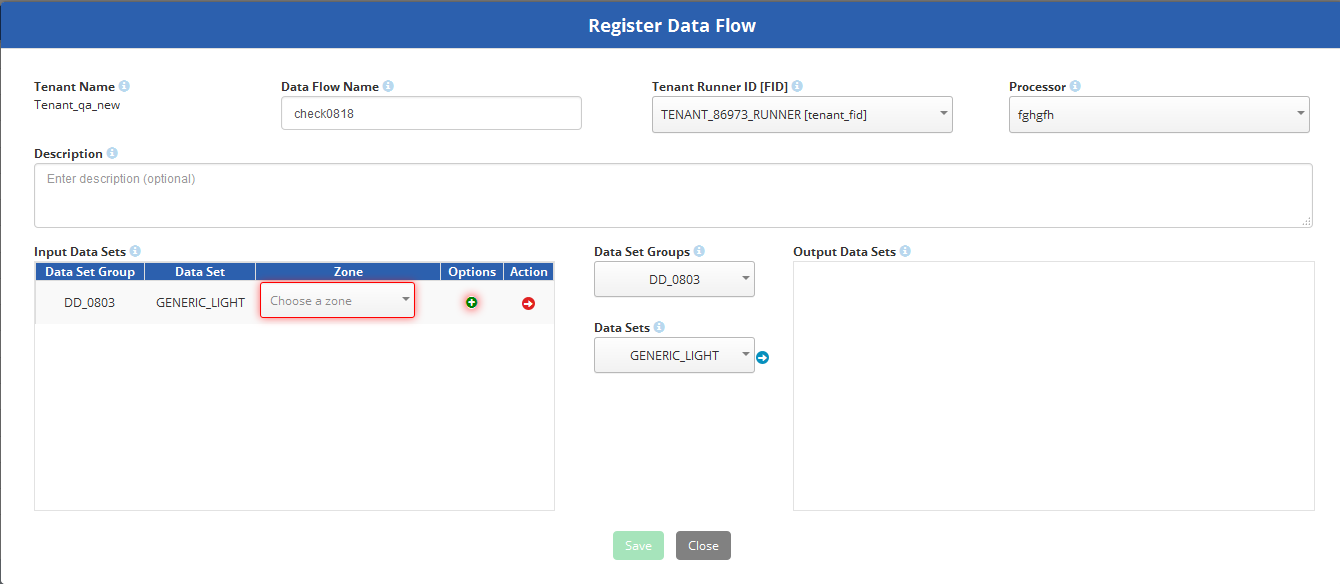
The ‘Tenant Name’ will appear for the tenant for which the user has logged in. A user will need to enter ‘Data Flow Name’, ‘Tenant Runner ID (FID)’, ‘Processor’ , these are the processors available under that tenant, ‘ Data Set Group’ and ‘Data Set’.

After the user enters the above field info, a left blue arrow  and a right blue arrow will appear next to the ‘Data Sets’ as shown below.

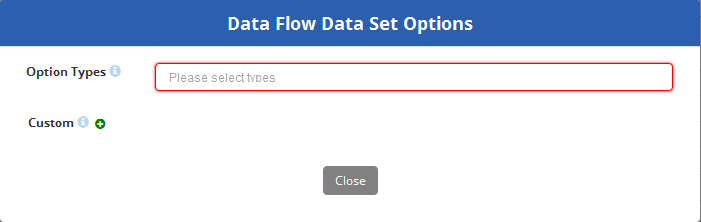


**Note**: For 1.1, a data set can be created without an input and output data set.

When a user wants to enter an ‘Input data set’ the left blue arrow  is selected and this activates the option to enter the input data sets as shown in the below figure



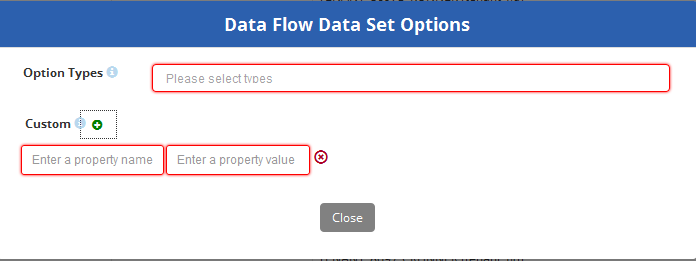
A user is required to select the ‘Zone’ from the drop down menu. A user is also required to enter the ‘Options’ fields next to the zones down drop menu. When ais clicked under the options, the ‘Data Flow Data Set Options’ screen pops up as shown in the below image.



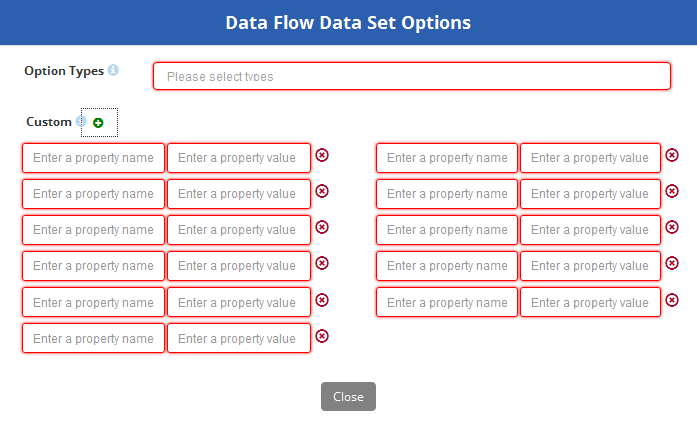
User is required to select the ‘Option Types’. Currently there are 3 different types of ‘Option Types’

* SQL
* File
* Table

The ‘Custom’ is not mandatory. It is used to enter ‘Property Name’ and ‘Property Value’

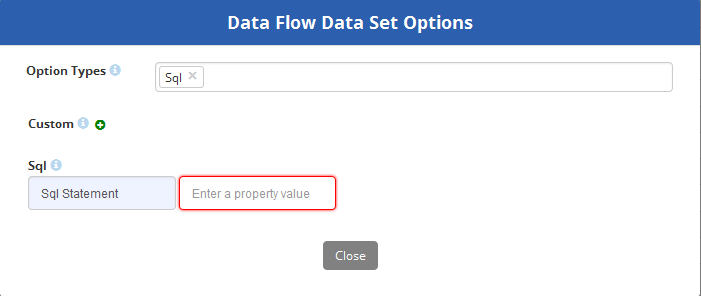


A user can enter multiple sets of values for ‘Property Name’ and ‘Property Value’ as shown in the below image.



# When Input ‘Option Types’= ‘Sql’

When a user selects ‘Sql’ as the ‘Option Type’ the option to enter the ‘Sql Statement’ appears as shown below.

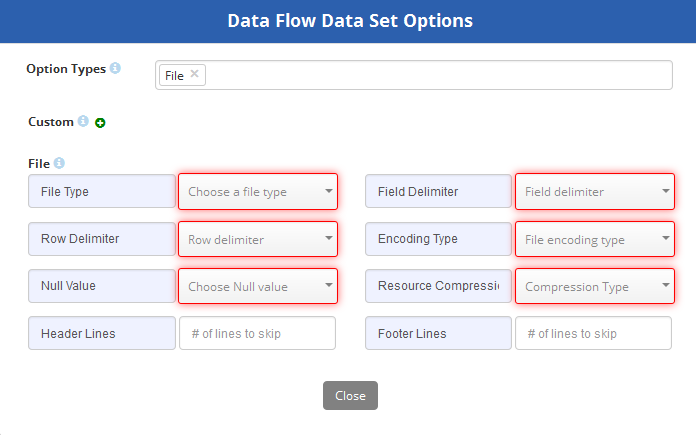


**Note:** Only Select commands are allowed with or without single or multiple joins. No DDL or DML Statements are allowed.

When the correct Sql statement is entered, and the  button is clicked it will take the user back to the ‘Register data Flow’ page where the user can proceed with selecting the output data set.

When Input ‘Option Types’= ‘File’

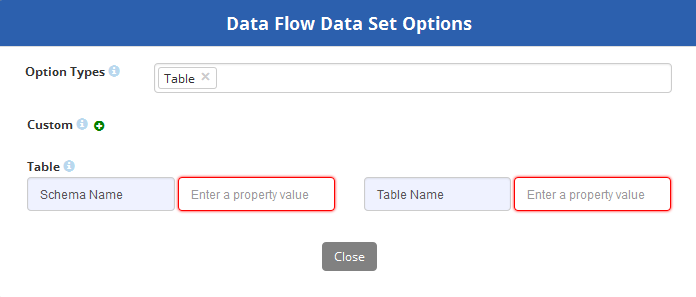
When a user selects ‘File’ as the ‘Option Types’ the following fields are enabled to enter the file options. All fields marked in red are mandatory



When the valid field values is entered of the fields marked in red, and the  button is clicked it will take the user back to the ‘Register data Flow’ page where the user can proceed with selecting the output data set.

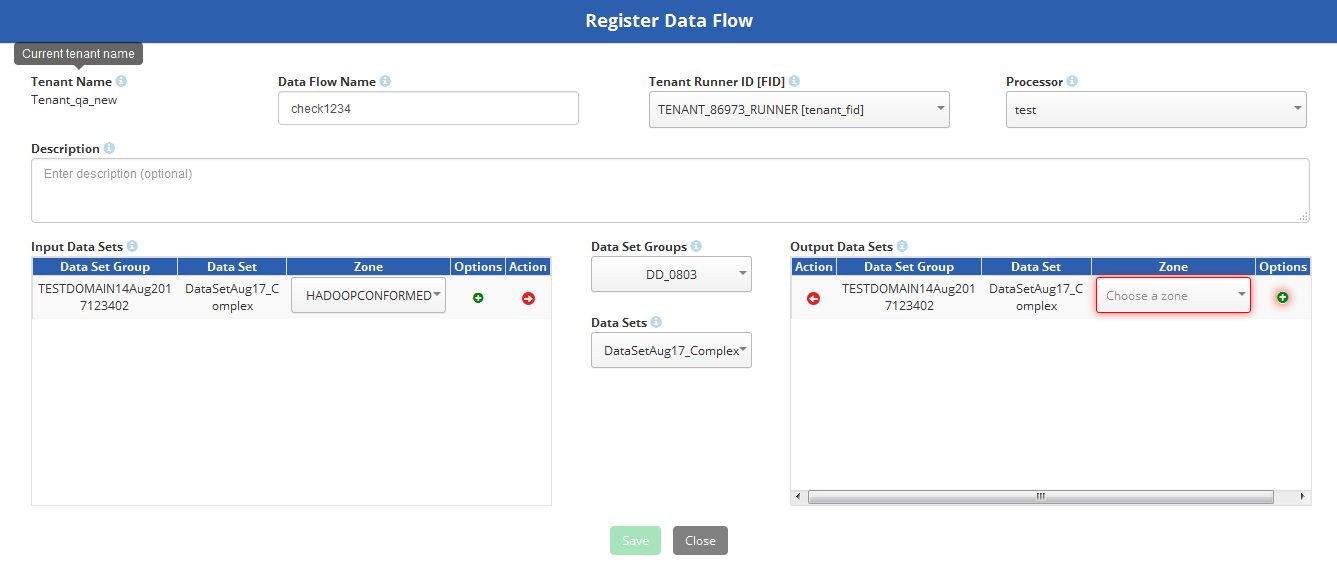
# When Input ‘Option Types’= ‘Table’

When a user selects ‘Table’ as the ‘Option Types’ the following fields are enabled to enter the file options. All fields marked in red are mandatory

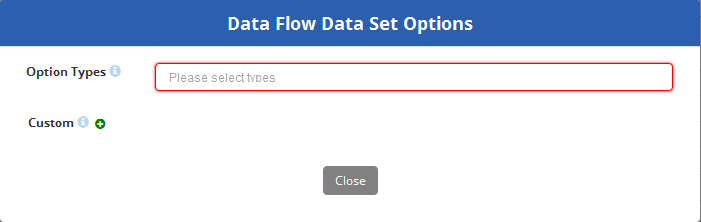


When the valid field values is entered of the fields marked in red, and the  button is clicked it will take the user back to the ‘Register data Flow’ page where the user can proceed with selecting the output data set.

Once an ‘Input Data Set’ has been selected, a user can select the ‘Output Data Set’ by clicking on the will make the fields to be entered for ‘Output Data Set’ visible as shown in the below image.



A user is required to select the Zones for the Output data set and then Options must be selected by clicking the  icon below the Options column. When ais clicked under the options, the ‘Data Flow Data Set Options’ screen pops up as shown in the below image.



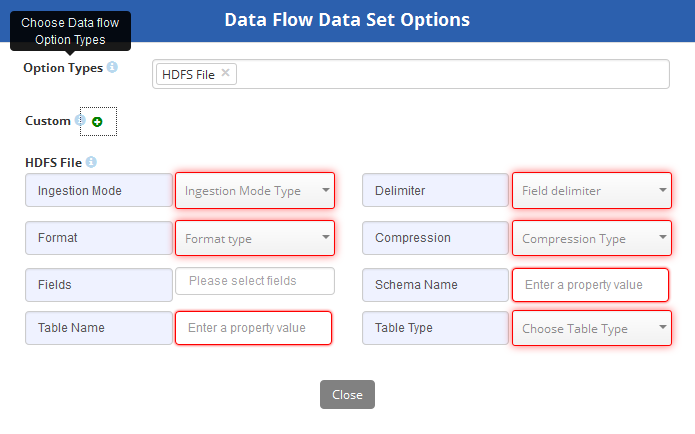
User is required to select the ‘Option Types’. Currently there are 2 different types of ‘Option Types’

* HDFS File
* Hive\_Impala

The ‘Custom’ is not mandatory. It is used to enter ‘Property Name’ and ‘Property Value’

# When Output ‘Option Types’= ‘HDFS File’

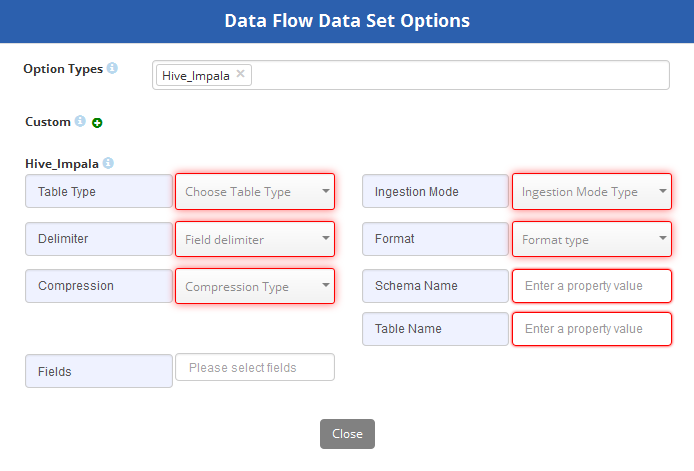
When a user selects ‘HDFS File’ as the ‘Option Types’ the following fields are enabled to enter the file options. All fields marked in red are mandatory



When the valid field values is entered of the fields marked in red, and the  button is clicked it will take the user back to the ‘Register data Flow’ page where the user can proceed with selecting the output data set.

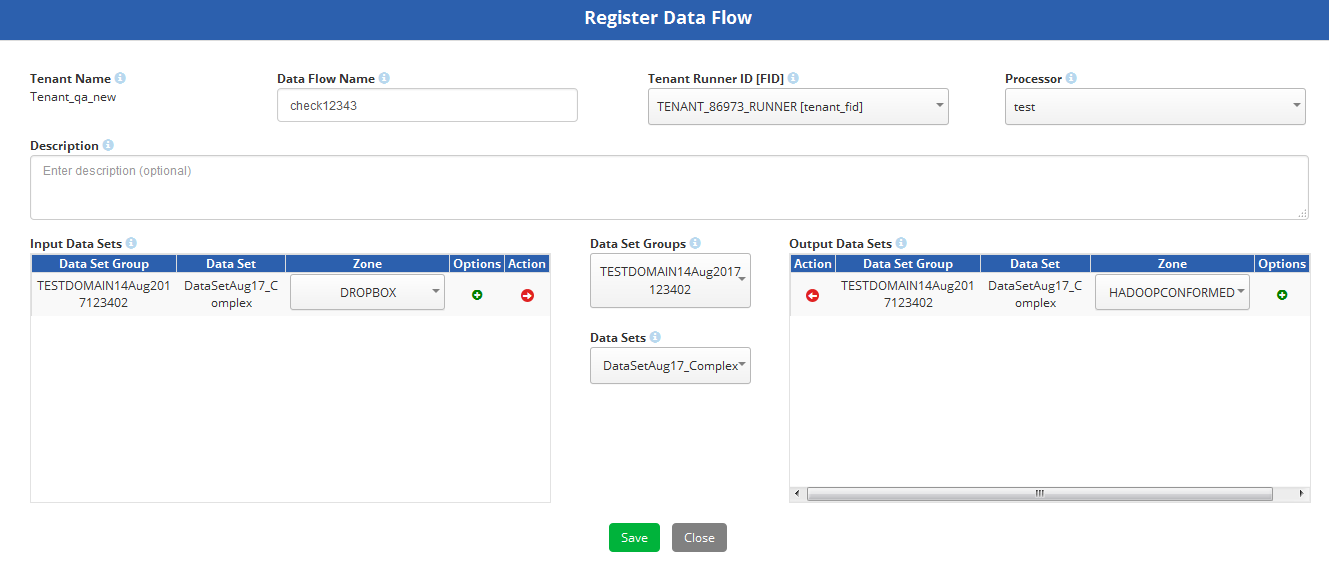
# When Output ‘Option Types’= ‘HDFS File’

When a user selects ‘Hive\_Impala’ as the ‘Option Types’ the following fields are enabled to enter the file options. All fields marked in red are mandatory



When the valid field values is entered of the fields marked in red, and the  button is clicked it will take the user back to the ‘Register data Flow’ page where the user can proceed with selecting the output data set.

When the ‘Input Data Set’ and ‘Output Data Set’ have been selected, the ‘Register Data Flow’ Page looks similar to the below image.

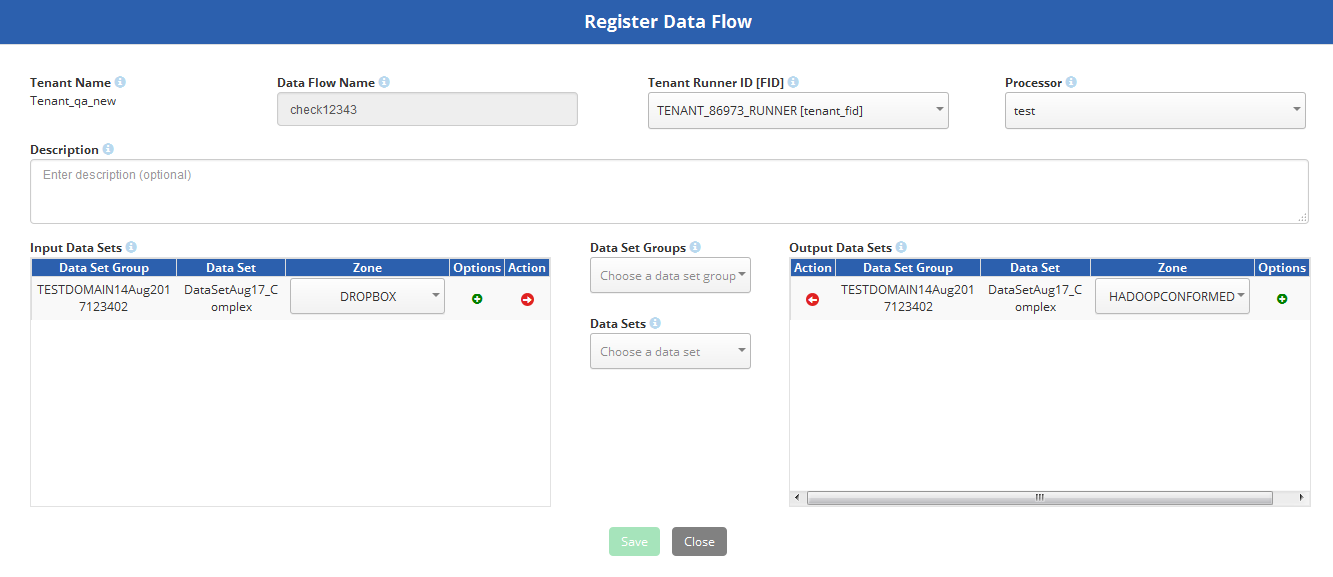


A user can then click on the  button to save the Data Flow. A message will appear confirming the data flow creation was successful as shown in the below image.



# Edit an existing Data Flow

To edit an existing Data Flow, click on the ‘Data Flow ID’ or ‘Data Flow Name’ in the data flow list page. A following pop up appears to edit the data flow.



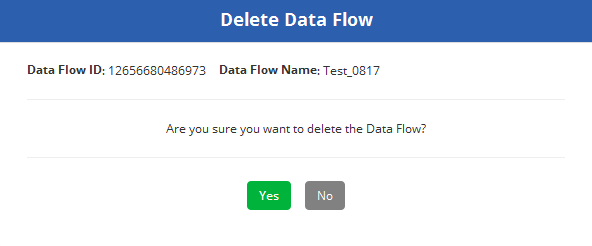
Except for the Data Flow Name, all the fields can be edited. A user can also edit the input and output data sets. Click on the  button to confirm the changes

# Delete an existing Data Flow

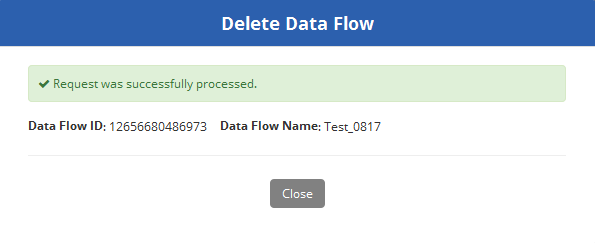
To delete an existing Data Flow, click on the  icon under the ‘Action’ column in the data flow list page. A following pop up appears to edit the data flow.



To delete the data flow, click on the ‘Delete’ option. Below screen is showing to confirm if the user wants to proceed with the deleting the data flow.



If the user clicks ‘Yes’, a confirmation page is shown confirming the delete was successful as shown in the below image.



If the user wants to promote the data flow from the lower to the higher environment, the ‘Promote’ feature is used. After the user selects the ‘Promote’ option for the data flow the user wants to promote it shows a pop up as shown in the below image

In the Data Flow list page, if the  icon is Red in color, this means that that data set has been already been promoted from lower to a higher environment.