

Contents

Datasets	3
View Defined Datasets	
Define a Dataset	
Execute a Dataset	(
Delete a Dataset	,

Datasets

A dataset is a specific view of data in PHEMI Central that is accessible to some limited set of users. The view provided by the dataset can be across multiple data sources.

The data elements available for building a dataset includes the derived data elements from all data sources across the data stored in PHEMI Central. You build a dataset by selecting data elements from among all those available. You can define any number of datasets from the fields available within PHEMI Central. Defining a dataset does not impact configuration for data sources or protections to data.

Datasets are typically created for research or exploratory data use, and the intended use for datasets may differ from the intended use of an individual data source. In particular, the governance and controls required for a dataset may differ from those specified its constituent data sources. For this reason, you can attach an access policy to a dataset independent of the access policies attached to the consistuent data sources.

Once a dataset is built, it can be executed by multiple users--even by users with different authorizations. A given user's actual view into the data in an executed dataset is restricted by his or her authorization. Thus, two users with different authorizations might execute the same dataset but end up with a different set of data.



Note: Only users with a Data Analyst role are permitted to build and execute datasets.

View Defined Datasets

View the list of defined datasets on the **Manage Datasets** page.

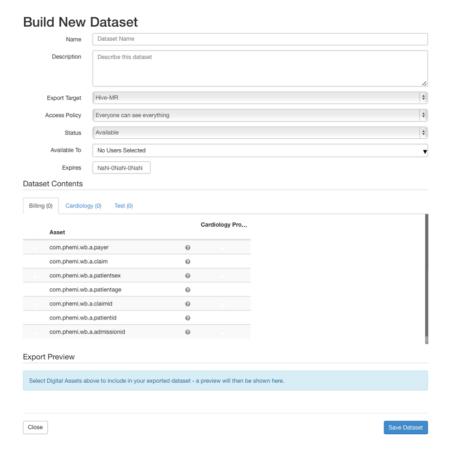
To see what datasets are available:

Open the Manage Datasets page, by clicking the Datasets icon in the left navigation bar.



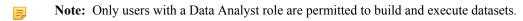
The **Manage Datasets** page opens, listing the datasets that have already been defined.

Datasets		
Select/Deselect All		
☐ BCCA - OMEGA Study	(not authorized)	
□ Cardiology Outcomes	Analysis(not authorized)	
COO Task Force - Hyp	othesis 1 (not authorized)	
☐ Test Cardiology Claims	(not authorized)	
USF Cardiology Exam	Die (not authorized)	
Execute Selected Datasets		



Define a Dataset

Define a new dataset on the Manage Datasets page.



To define a dataset:

1.

Open the **Manage Datasets** page, by clicking the **Datasets** icon in the left navigation bar.

The **Manage Datasets** page opens, listing the datasets that have already been defined.



2. Click the **Define New Dataset** button.

The Build New Dataset screen opens.

Name	Dataset Name				
Description	Describe this dataset				
Export Target	Hive-MR				
Access Policy	Everyone can see everything	ng			
Status	Available				
Available To	No Users Selected				
Expires	NaN-0NaN-0NaN				
ataset Contents	11211 011211 011211				
Billing (0) Cardiolo	gy (0) Test (0)	c	ardiology Pro		
Asset	0.0000				
com.phemi.wb		0			
com.phemi.wb		0			
com.phemi.wb	.a.patientage	0			
	.a.claimid	0			
com.phemi.wb		0			
com.phemi.wb	.a.patientid				
		0			
com.phemi.wb		0			
com.phemi.wb		0			
com.phemi.wb com.phemi.wb xport Preview			review will then be show	wn here.	
com.phemi.wb com.phemi.wb xport Preview	.a.admissionid		eview will then be show	wn here.	

Execute a Dataset



Note: Only users with a Data Analyst role are permitted to build and execute datasets, and Data Analyst users may only build and execute datasets for which they have a requisite permissions.

A subset of the elements produced by Data Processing Functions are selected to define a dataset that will be materialized in HANA. The chosen Access Policies determine the data representation (e.g. identified or de-identified) that is persisted in HANA.

Delete a Dataset