

Datasets

Contents

- Datasets..... 3**
 - View Defined Datasets..... 3
 - Define a Dataset..... 4
 - Execute a Dataset..... 6
 - Delete a Dataset..... 7

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 constituent 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.

Manage Datasets

Datasets

☐ Select/Deselect All

☐ BCCA - OMEGA Study(not authorized)

☐ Cardiology Outcomes Analysis(not authorized)

☐ COO Task Force - Hypothesis 1(not authorized)

☐ Test Cardiology Claims(not authorized)

☐ USF Cardiology Example(not authorized)

Build New Dataset

Name

Dataset Name

Description

Describe this dataset

Export Target

Hive-MR

Access Policy

Everyone can see everything

Status

Available

Available To

No Users Selected

Expires

NaN-0NaN-0NaN

Dataset Contents

Billing (0)

Cardiology (0)

Test (0)

Asset

Cardiology Pro...

com.phemi.wb.a.payer		
com.phemi.wb.a.claim		
com.phemi.wb.a.patientsex		
com.phemi.wb.a.patientage		
com.phemi.wb.a.claimid		
com.phemi.wb.a.patientid		
com.phemi.wb.a.admissionid		

Export Preview

Select Digital Assets above to include in your exported dataset - a preview will then be shown here.

Close

Save Dataset

Define a Dataset

Define a new dataset on the **Manage Datasets** page.



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

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.

Manage Datasets

Datasets

☐ Select/Deselect All

☐ BCCA - OMEGA Study(not authorized)

☐ Cardiology Outcomes Analysis(not authorized)

☐ COO Task Force - Hypothesis 1(not authorized)

☐ Test Cardiology Claims(not authorized)

☐ USF Cardiology Example(not authorized)

Execute Selected Datasets

Define New Dataset

2. Click the **Define New Dataset** button.
The Build New Dataset screen opens.

Build New Dataset

Name

Description

Export Target

Access Policy

Status

Available To

Expires

Dataset Contents

Asset	Cardiology Pro...
<input type="checkbox"/> com.phemi.wb.a.payer	<input type="checkbox"/>
<input type="checkbox"/> com.phemi.wb.a.claim	<input type="checkbox"/>
<input type="checkbox"/> com.phemi.wb.a.patientsex	<input type="checkbox"/>
<input type="checkbox"/> com.phemi.wb.a.patientage	<input type="checkbox"/>
<input type="checkbox"/> com.phemi.wb.a.claimid	<input type="checkbox"/>
<input type="checkbox"/> com.phemi.wb.a.patientid	<input type="checkbox"/>
<input type="checkbox"/> com.phemi.wb.a.admissionid	<input type="checkbox"/>

Export Preview

Select Digital Assets above to include in your exported dataset - a preview will then be shown 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
