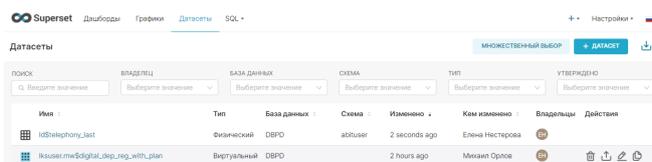


Содержание

Датасеты (Datasets)	2
Создание датасета на основе таблицы	3
Настройка датасета	3
Источник	3
Меры	4
Столбцы	4
Вычисляемые столбцы	5
Настройки	5
Изменение датасета на произвольный SQL-запрос	6
Удаление датасета	7

Датасеты (Datasets)

Датасет (источник данных) – описание набора данных, которые хранятся в базе данных. Сами данные не хранятся в Superset, а только загружаются при запросе. Раздел с перечнем всех датасетов в разделе **Датасеты**.



Справа располагаются кнопки для **множественного выбора** из существующих датасетов, **добавления датасета** из базы данных и из файла. Ниже располагаются окно **поиска** и окна для **фильтрации** загруженных датасетов по таким полям, как:

- **Владелец** датасета;
- Используемая **база данных**;
- Используемая **схема**;
- **Тип** датасета: физический  или виртуальный  ;
- **Утверждено**- сертифицирован ли датасет.

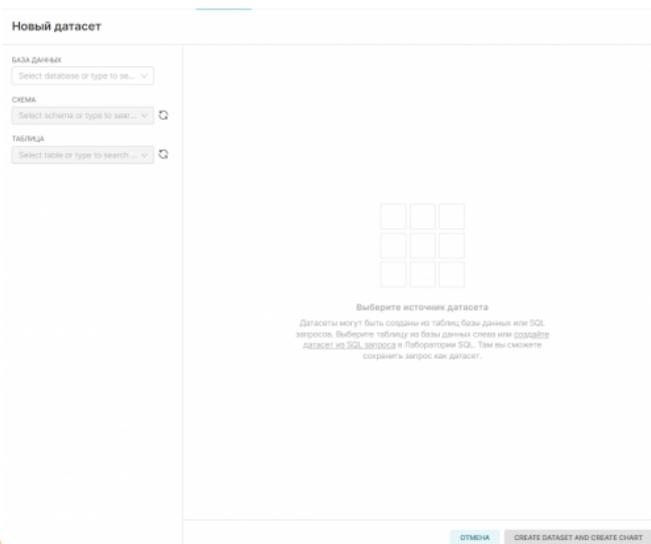
Для датасетов отображаются следующие атрибуты:

- **Имя** – название датасета;
- **Тип** – физический или виртуальный;
- **База данных** – используемая база данных;
- **Схема** – используемая схема;
- **Изменено** – когда последний раз был изменен датасет;
- **Кем изменено** – кем в последний раз был изменен датасет;
- **Владельцы** – пользователи, которые могут изменять датасет;
- **Действия** – доступные действия, которые можно выполнить базой данных.
 - Удалить;
 - Экспортировать;
 - Редактировать
 - Дублировать (только для виртуальных датасетов).

Создание датасета на основе таблицы

В разделе Датасеты справа нажимаем на кнопку + **ДАТАСЕТ**. В диалоговом окне выбираем базу данных, схему и таблицу.

После выбора таблицы справа появятся столбцы таблицы и их тип. Нажмите внизу на кнопку **Create dataset and create chart**. Такой созданный датасет будет называться физическим, так как данные датасета физически хранятся в указанной таблице. Физический датасет в списке выделяется черно-белым .



Новый датасет

БАЗА ДАННЫХ
Select database or type to add...

СХЕМА
Select schema or type to add...

ТАБЛИЦА
Select table or type to select...

Выберите источник датасета.
Датасеты могут быть созданы из таблиц, базы данных или SQL-запросов. Выберите таблицу из базы данных слева или создайте датасет из SQL-запроса в Лаборатории SQL. Там вы сможете сохранить запрос как датасет.

ОТМЕНА CREATE DATASET AND CREATE CHART



Так как при создании можно создать только физический датасет, у которого нельзя изменить название (т.к. название датасета= название таблицы), то два физических датасета с одинаковыми именами таблиц не могут быть созданы (даже если они в разных схемах и базах данных).

Если нужно создать несколько датасетов ссылающихся на одну таблицу, то сделайте два виртуальных датасета, где вручную пропишите выборку SQL-запросом к одной и той же таблице (см. подробнее следующий раздел).

Настройка датасета

Каждый датасет после создания можно изменить. У выбранного датасета в столбце **Действия** выбираем значок карандаша.

Источник

На вкладке **Источник** можно просмотреть текущий источник данных. Редактирование на вкладке защищено от случайного изменения.

Меры

На вкладке **Меры** можно настроить показатели (метрики) датасета. Метрики используются для выполнения операций с набором данных (т.е. применяются более чем к одной строке), например, с агрегатными функциями в Pivot-таблицах.

Чтобы добавить показатель нажмите кнопку **+ ДОБАВИТЬ ЗАПИСЬ**. В новой строке в поле **Metric Key** введите название метрики, которое будет отображаться в списке. В поле **Метка** введите название, которое будет отображаться в конструкторе графика.

Выражение SQL должно содержать агрегатные или иные функции. Корректность выражения не проверяется в данном окне. Слева по нажатию на ▶ у каждой метрики можно раскрыть область дополнительных параметров, ввести описание, формат вывода данных, информацию об утверждении данных и предупреждающее сообщение (появляется в поле выбора метрики в конструкторе графиков), параметры валюты и др. Чтобы **удалить** строку с метрикой, нажмите справа значок корзины.

Столбец	Тип данных	Содержит дату/время	Дата и время по умолчанию	Фильтруемый	Является измерением
unit	DATE	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
pk	УНИКАЛЬНЫЙ СЧЕТ	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
type_of_name	УНИКАЛЬНЫЙ СЧЕТ	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
client	УНИКАЛЬНЫЙ СЧЕТ	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
employee	УНИКАЛЬНЫЙ СЧЕТ	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
window	УНИКАЛЬНЫЙ СЧЕТ	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Столбцы

Вкладка **Столбцы** содержит описание столбцов, определенное таблицей или SQL-запросом источника. **Тип данных** содержит тип данных столбца (например, integer, float, variable character, timestamp и т.д.), это поле не редактируемое.

Столбец **Содержит дату/время** указывает, что данные содержать дату или время в каком-либо формате (date, time, timestamp и т.д.).

Дата и время по умолчанию указывает на столбец, который будет выбран по умолчанию, как дата и время.

Флаг **Фильтруемый** разрешает включать столбец датасета в фильтр на графиках/дашбордах.

Флаг **Является измерением** указывает, можно ли по данному столбцу применять «GROUP BY» в запросе, то есть можно ли группировать в pivot-таблицах.

Значок корзины для **удаления** столбца.

У каждого столбца датасета можно

настроить дополнительные параметры, нажав на стрелку ▶ :

- **Метка** – отображаемое название столбца в графиках (там можно переопределить);
- **Описание** – краткое описание столбца, появляется при наведении курсора;
- **Формат даты/времени** – если у столбца выставлен флаг **Содержит дату/время**, то дата и время будут преобразовываться для вывода по заданному формату. Формат задается в соответствии с форматом Python (например, %m-%d-%Y= 02-27-2020);
- **Кем утверждено** – пользователь(ли), утвердивший данный показатель;
- **Детали утверждения.**

Вычисляемые столбцы

Вкладка **Вычисляемые столбцы** содержит столбцы, которые автор датасета может сам задать. Эти столбцы физически не хранятся ни в базе данных, ни в superset, они вычисляются “на лету” при выполнении SQL-запроса. Настройки у вычисляемых столбцов такие же, как и на вкладке **Столбцы**, однако добавляются поля **Выражение SQL** и **Тип данных**.

- Например, из двух значений физических столбцов *department* и *owner* нужно получить одну строку с помощью их конкатенации, тогда расчётный столбец будет задан следующим образом, в Выражение SQL прописываем соответствующие операторы, Метка будет определять понятное название на русском языке, а в поле Таблица данных (прим. тип данных) выбираем типа string.

Настройки

На вкладке **Настройки** можно изменить описание датасета, время обновления данных, часовой пояс для смещения времени относительно UTC, владельцев (пользователей, которые могут изменять датасет).

После настройки датасета нажмите кнопку **Сохранить** для применения изменений.

Изменение датасета на произвольный SQL-запрос

Создать датасет сразу с произвольным SQL-запросом нельзя (или воспользуйтесь [Лабораторией SQL](#)). Можно изменить таблицу на произвольный запрос только в уже существующем датасете.

Датасет основанный на произвольном SQL-запросе называется **виртуальным**, потому что с помощью запроса данные датасета могут быть получены, как и обычной выборкой реальных данных из физической таблицы, так и с помощью разных инструментов (например, агрегатные, аналитические или оконные функции) из разных таблиц разных схем. Виртуальный датасет в списке выделяется синим .

Для смены типа датасета на виртуальный нажмите на значок карандаша в действиях. В диалоговом окне настроек на вкладке **Источник** заблокировано изменение в целях случайного изменения. Нажмите на , чтобы разблокировать редактирование полей и выберите ниже **Виртуальный (SQL)**. В поле SQL можно ввести любой SQL-запрос. При вводе запроса никаких проверок на корректность синтаксиса и выборки данных не производится.

После ввода запроса нажимаем снизу кнопку **Сохранить** и подтверждаем применение изменений. Появится сообщение

 Источник данных, похоже, был удален , что значит удаление старого датасета и создание нового.

 Нажмите на замок, чтобы предотвратить будущие изменения.

Физический (таблица или представление) Виртуальный (SQL)

VIRTUAL

БАЗА ДАННЫХ
oracle DBPD

СХЕМА
abituser

НАИМЕНОВАНИЕ ДАТАСЕТА
Id\$stat_orig

SQL 

```

1 select vu, cnt, dt,
2       case when vu2='НГТУ' then 'Новосибирский государственный технический университет' (НГТУ) then 'НГТУ'
3       else 'Другие ВУЗы' end is_ngtu
4 from abituser.Id$stat_orig Id

```



По сути, предыдущий датасет удаляется и создается новый, поэтому еще раз подумайте, возможно не стоит трогать уже существующий, а лучше создать новый.

Чтобы обновилась таблица в соответствии с результатом запроса, снова из списка датасетов заходим в окно редактирования, сверху выбираем вкладку Столбцы. По кнопке **Синхронизировать столбцы из источника** столбцы ниже обновятся, добавятся новые (если были в запросе).

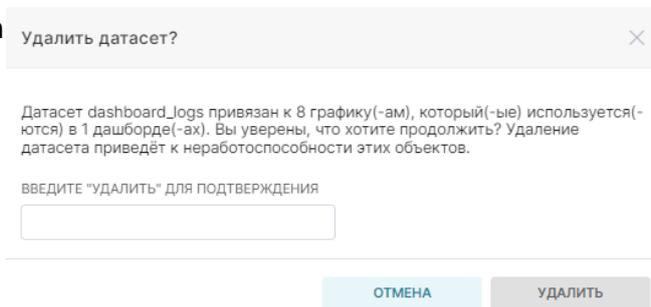
Также появится сообщение в углу об обновлении столбцов датасета, что говорит о правильном выполнении SQL-запроса.



Если SQL-запрос синтаксически неправильно написан или произошла ошибка при синхронизации, будет показано соответствующее сообщение с описанием ошибки.

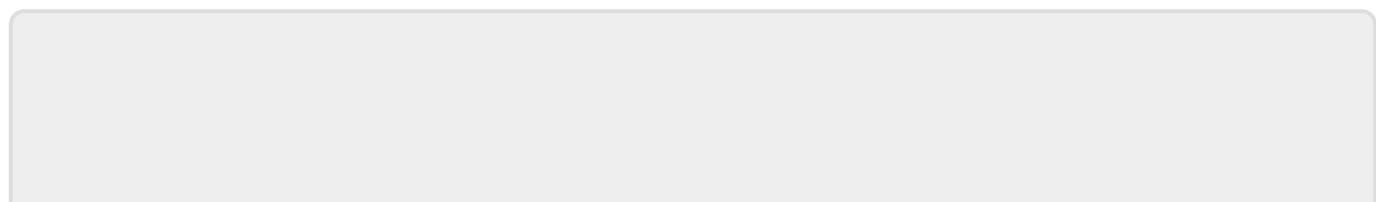
Удаление датасета

В списке датасетов наведите курсор мыши на строку нужного датасета в области столбца **Действия**. Появятся значки взаимодействия с датасетом. Нажмите на значок **корзины**. В появившемся диалоговом окне будет указано к какому количеству графиков и дашбордов привязан выбранный датасет. После удаления датасета графики станут нерабочими.



Для удаления введите в строке удалить и нажмите на кнопку **Удалить**, которая станет доступна после ввода.

 Восстановить удаленный датасет будет невозможно, только создать/добавить заново!



From:

<https://kb.nstu.ru/> - База знаний НГТУ НЭТИ

Permanent link:

https://kb.nstu.ru/superset:work_in_superset:datasets:datasets

Last update: **2023/11/15 15:08**

