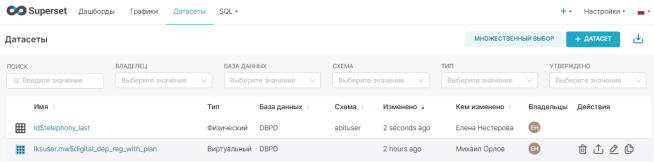
Содержание

Датасеты (Datasets) 2
Создание датасета на основе таблицы 3
Настройка датасета 3
Источник 3
Меры 4
Столбцы 4
Вычисляемые столбцы 5
Настройки 5
Работа с датасетом 6

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**.

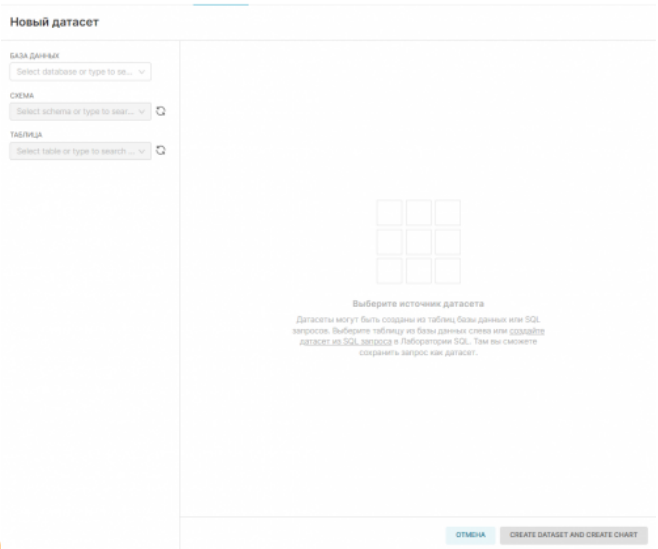
Справа располагаются кнопки для **множественного выбора** из существующих датасетов, **добавления датасета** из базы данных и из файла. Ниже располагаются окно **поиска** и окна для **фильтрации** загруженных датасетов по таким полям, как:

- **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 выполнитьс 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 .

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).

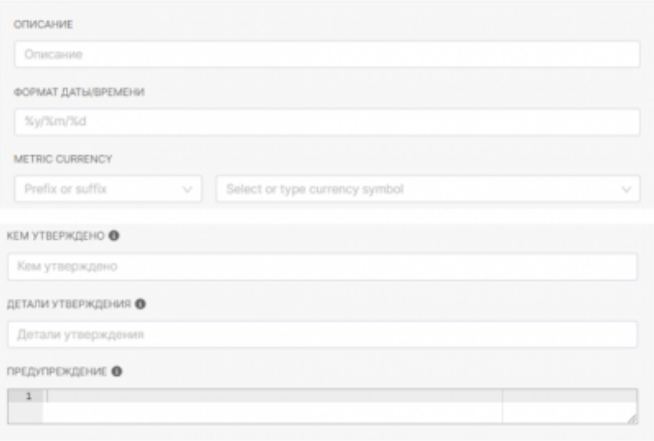
Настройка датасета

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

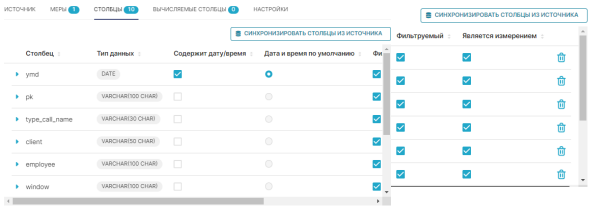
Источник

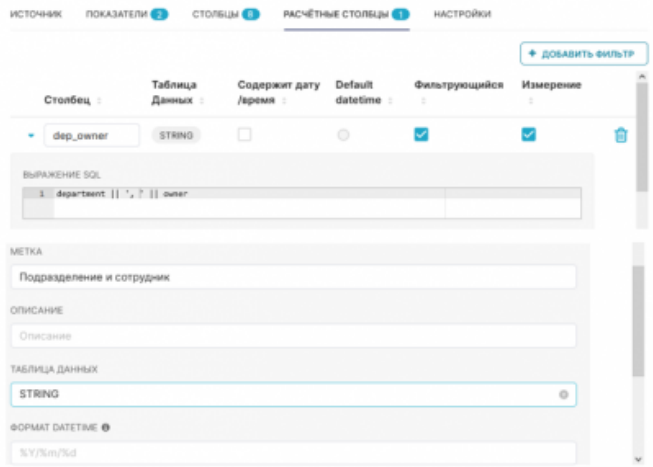
На вкладке **Источник** можно просмотреть текущий источник данных. Редактирование на вкладке защищено от случайного изменения.

Меры

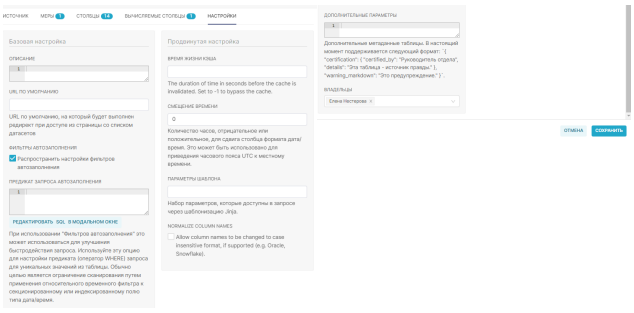


На вкладке **Меры** можно настроить показатели (метрики) датасета. Метрики используются для выполнения операций с набором данных (т.е. применяются более чем к одной строке), например, с агрегатными функциями в Pivot-таблицах.
Чтобы добавить показатель нажмите кнопку **+ ДОБАВИТЬ ЗАПИСЬ**. В новой строке 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.





Столбцы



Вкладка **Столбцы** содержит описание столбцов, определенные 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 дополнительный 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 настроки датасета нажмите кропку
Сохранить для применения изменений.

Работа с датасетом

⇒ Создание датасета на основе таблицы
⇒ Настройки датасета
⇒ Изменение датасета на произвольный SQL-запрос
⇒ Удаление датасета

From:
https://kb.nstu.ru/ - **База знаний НГТУ НЭТИ**

Permanent link:
https://kb.nstu.ru/superset:work_in_superset:datasets:datasets?rev=1700032105

Last update: **2023/11/15 14:08**