
Регистрация нового датасета

В данном разделе описан процесс регистрации датасета из источника, подключенного к системе. Подробнее про источники данных см. раздел "Источники данных".


Важно

На данный момент существует **ограничение** на объем датасета - **1 Гб**. Система не тестировалась на работу с файлами большего размера. Рекомендуемый размер файла - до **100 Мб**.

Переход осуществляется через вкладку **Данные** в верхней панели навигации. В открывшемся окне **Каталога данных** нажмите кнопку "Добавить".

Появится форма регистрации нового датасета:

Создать данные



Источник данных не выбран

[Выбрать](#)

* Название

Описание

Метки


[Отмена](#) [Создать](#)

В первую очередь необходимо выбрать источник данных. При нажатии на кнопку "Выбрать" появится возможность выбрать нужный источник.


После выбора источника форма создания датасета обновится в соответствии с типом выбранного источника. Если выбран источник типа S3, то форма будет содержать

поле для загрузки файла:

Создать данные


s3-prod 

Файл



Щёлкните или перетащите файл в эту область, чтобы загрузить

Поддерживает загрузку одного файла

* Скрипт для запроса данных 

Список таблиц

Столбец, содержащий дату наблюдения

* Название

Описание

Метки

Параметры регистрации

- **Источник данных** - необходимо выбрать из списка доступных источников данных
- **Файл** - поле для загрузки файла на S3. Доступно только для источников данных S3. Можно щелкнуть или перетащить файл в указанную область.

- **Скрипт для запроса данных** (обязательное поле) - скрипт на том диалекте SQL, который воспринимается источником данных.



Подсказка для S3

Для источников типа S3 будет автоматически предложена функция для чтения всего датасета, например: `data/predictions.csv`

- **Столбец, содержащий дату наблюдения** (необязательное поле) - название столбца, который содержит отчетную дату для данных. Выбирается после составления скрипта и нажатия кнопки "Выполнить".
- **Название** (обязательное поле) - название датасета.
- **Описание** (необязательное поле) - описание датасета.
- **Метки** (необязательное поле) - метки (теги) датасета.

Для просмотра списка доступных таблиц можно воспользоваться кнопкой "Список таблиц".

После заполнения всех необходимых данных нажмите кнопку "Создать". Для отмены операции нажмите "Отмена".

Выбор существующих файлов в S3

Если файл с данными уже был загружен в S3 хранилище, нет необходимости загружать его повторно. Вы можете использовать существующие файлы:

1. Нажмите кнопку "Список таблиц"
2. В появившемся списке выберите нужный файл:

Название	Скрипт для запроса данных
data/100mb30col_2024_11_21_15_53_46_...	data/100mb30col_2024_11_21_15_53_46_670.csv 🔗
data/100mb30col_2024_11_21_15_54_38_...	data/100mb30col_2024_11_21_15_54_38_699.csv 🔗
data/1800mb30col.csv	data/1800mb30col.csv 🔗
data/1gb.csv	data/1gb.csv 🔗
data/TEST_2024_9_24_11_57_59_678.csv	data/TEST_2024_9_24_11_57_59_678.csv 🔗
data/TEST_2024_9_24_12_43_20_897.csv	data/TEST_2024_9_24_12_43_20_897.csv 🔗
data/TEST_2024_9_24_13_16_44_767.csv	data/TEST_2024_9_24_13_16_44_767.csv 🔗

[Заккрыть](#)

- В колонке **Название** отображается имя файла в хранилище S3
- В колонке **Скрипт для запроса данных** указан путь к файлу для использования в SQL-запросе. Скопируйте его, чтобы в дальнейшем его вставить в поле основного меню **Скрипт для запроса данных**.