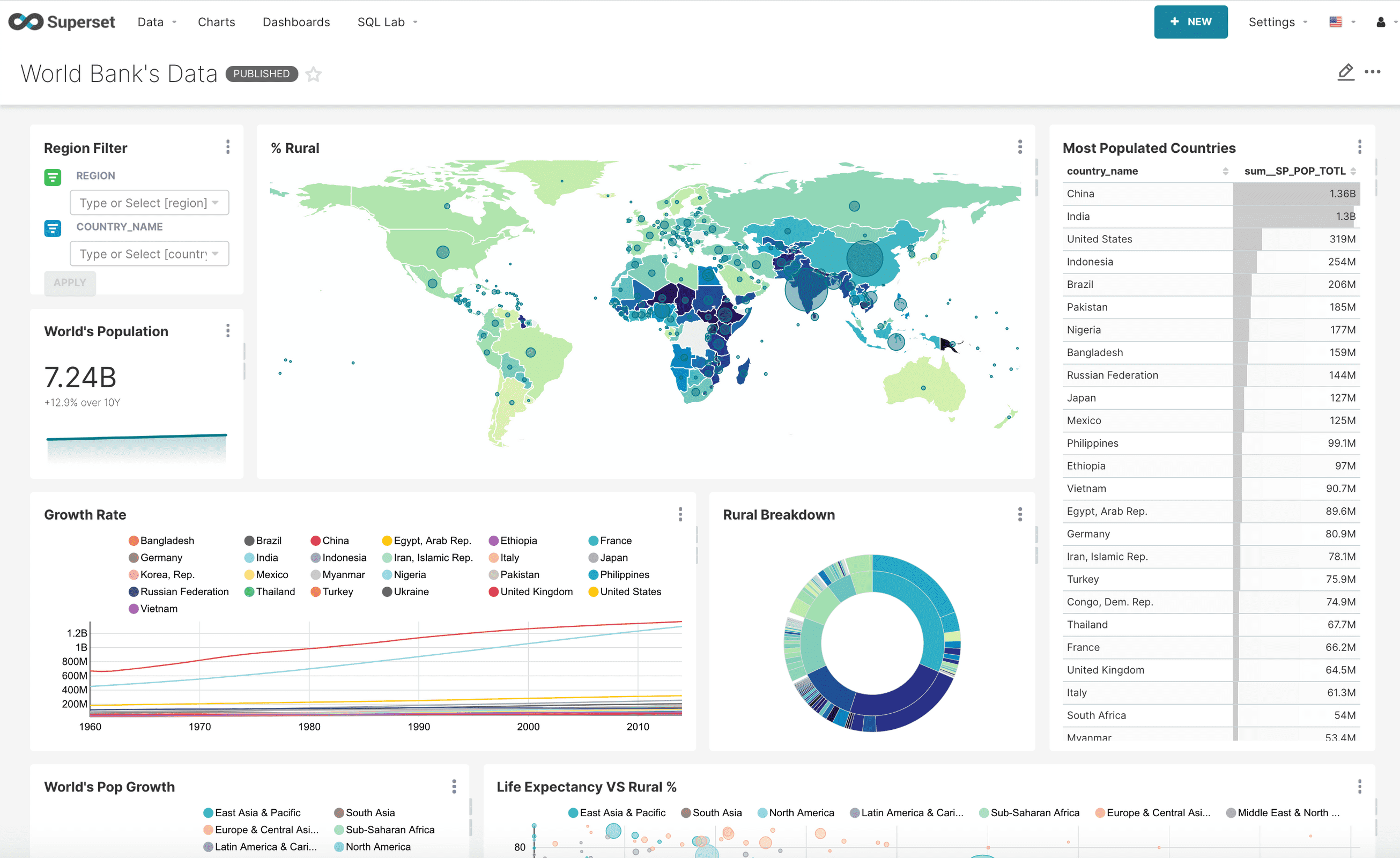
# 第1章 Superset入门

## 1.1 Superset概述

Apache Superset是一个现代的数据探索和可视化平台。它功能强大且十分易用，可对接各种数据源，包括很多现代的大数据分析引擎，拥有丰富的图表展示形式，并且支持自定义仪表盘。



## 1.2 环境说明

本课程使用的服务器操作系统为CentOS 7，Superset对接的数据源为MySQL数据库。

# 第2章 Superset安装

Superset官网地址：http://superset.apache.org/

## 2.1 安装Python环境

Superset是由Python语言编写的Web应用，要求Python3.7以上的环境。

### 2.1.1 安装Miniconda

conda是一个开源的包、环境管理器，可以用于在同一个机器上安装不同Python版本的软件包及其依赖，并能够在不同的Python环境之间切换，Anaconda包括Conda、Python以及一大堆安装好的工具包，比如：numpy、pandas等，Miniconda包括Conda、Python。

此处，我们不需要如此多的工具包，故选择MiniConda。

**1）下载Miniconda（Python3版本）**

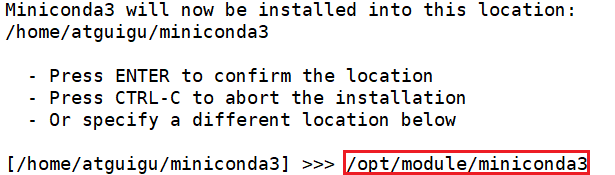
下载地址：<https://repo.anaconda.com/miniconda/Miniconda3-latest-Linux-x86_64.sh>

**2）安装Miniconda**

（1）执行以下命令进行安装，并按照提示操作，直到安装完成。

[test1@master lib]$ bash Miniconda3-latest-Linux-x86\_64.sh

（2）在安装过程中，出现以下提示时，可以指定安装路径 /opt/module/miniconda3



（3）出现以下字样，即为安装完成



**3）加载环境变量配置文件，使之生效**

[test1@master lib]$ source ~/.bashrc

**4）取消激活base环境**

Miniconda安装完成后，每次打开终端都会激活其默认的base环境，我们可通过以下命令，禁止激活默认base环境。

[test1@master lib]$ conda config --set auto\_activate\_base false

### 2.1.2 创建Python3.8环境

**1）配置conda国内镜像（此处可不配置，国外镜像比清华镜像更快）**

(base) [test1@master ~]$ conda config --add channels https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/free

(base) [test1@master ~]$ conda config --add channels https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/main

(base) [test1@master ~]$ conda config --set show\_channel\_urls yes

**（1）查看conda的镜像channel配置**

[test1@master ~]$ conda config --show channels

- https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/main

- https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/free

- defaults

**（2）删除清华镜像,继续用默认的**

[test1@master ~]$ conda config --remove-key channels

[test1@master ~]$ conda config --show channels

- defaults

**2）创建Python3.8环境**

(base) [test1@master ~]$ conda create --name superset python=3.8.16

* **说明：conda环境管理常用命令**
* **创建环境：**conda create -n env\_name
* **查看所有环境：**conda info --envs
* **删除一个环境：**conda remove -n env\_name --all

**3）激活superset环境**

(base) [test1@master ~]$ conda activate superset

激活后效果如下图所示



**说明：退出当前环境**

(superset) [test1@master ~]$ conda deactivate

**4）执行python -V命令查看python版本**

(superset) [test1@master ~]$ python -V

**Python 3.8.16**

## 2.2 Superset部署

### 2.2.1 安装依赖

安装Superset之前，需安装以下所需依赖。

(superset) [test1@master ~]$ sudo yum install -y gcc gcc-c++ libffi-devel python-devel python-pip python-wheel python-setuptools openssl-devel cyrus-sasl-devel openldap-devel

### 2.2.2 安装Superset

**1）更新pip**

(superset) [test1@master ~]$ pip install --upgrade pip -i https://pypi.douban.com/simple/

**说明：**pip是python的包管理工具，可以和centos中的yum类比

**2）安装Supetset**

（1）上传base.txt文件至任意路径

(superset) [test1@master ~]$ ll ~

总用量 12

-rw-r--r--. 1 test1 test1 5795 2月 3 15:27 **base.txt**

drwxrwxr-x. 2 test1 test1 4096 2月 6 11:42 bin

该文件可用于指定superset依赖组件及版本，下载地址及内容如下。

<https://raw.githubusercontent.com/apache/superset/2.0.0/requirements/base.txt>



（2）在base.txt所在目录下执行如下命令，安装SuperSet

(superset) [test1@master ~]$ pip install apache-superset==2.0.0 -i https://pypi.tuna.tsinghua.edu.cn/simple -r base.txt

**说明：**

-i的作用是指定镜像，这里选择国内镜像

-r的作用是指定superset依赖组件及相应版本，指向base.txt文件即可

### 2.2.3 配置Superset元数据库

Superset的元数据支持MySQL、PostgreSQL，此处采用MySQL。

**1）在MySQL中创建superset元数据库**

mysql> CREATE DATABASE superset DEFAULT CHARACTER SET utf8 DEFAULT COLLATE utf8\_general\_ci;

**2）创建superset用户**

mysql> create user superset@'%' identified WITH mysql\_native\_password BY 'superset';

mysql> grant all privileges on \*.\* to superset@'%' with grant option;

mysql> flush privileges;

**3）修改superset配置文件**

(superset) [test1@master ~]$ vim /opt/module/miniconda3/envs/superset/lib/python3.8/site-packages/superset/config.py

修改内容如下：（184、185行）

**#** SQLALCHEMY\_DATABASE\_URI = "sqlite:///" + os.path.join(DATA\_DIR, "superset.db")

SQLALCHEMY\_DATABASE\_URI = 'mysql://**superset**:**superset**@**master**:3306/**superset**?charset=utf8'

**4）安装python msyql驱动**

(superset) [test1@master ~]$ conda install mysqlclient

**5）初始化superset元数据**

(superset) [test1@master ~]$ export FLASK\_APP=superset

(superset) [test1@master ~]$ superset db upgrade

### 2.2.4 SupersetSet初始化

**1）创建管理员用户**

(superset) [test1@master ~]$ superset fab create-admin

**2）初始化superset**

(superset) [test1@master ~]$ superset init

### 2.2.5 启动Supterset

**1）安装gunicorn**

(superset) [test1@master ~]$ pip install gunicorn -i https://pypi.douban.com/simple/

**说明：**gunicorn是一个Python Web Server，可以和java中的TomCat类比

**2）启动Superset**

**（1）确保当前conda环境为superset，及下图所示**



**（2）启动**

(superset) [test1@master ~]$ gunicorn --workers 5 --timeout 120 --bind master:8787 "superset.app:create\_app()" --daemon

**说明：**

**--workers：指定进程个数**

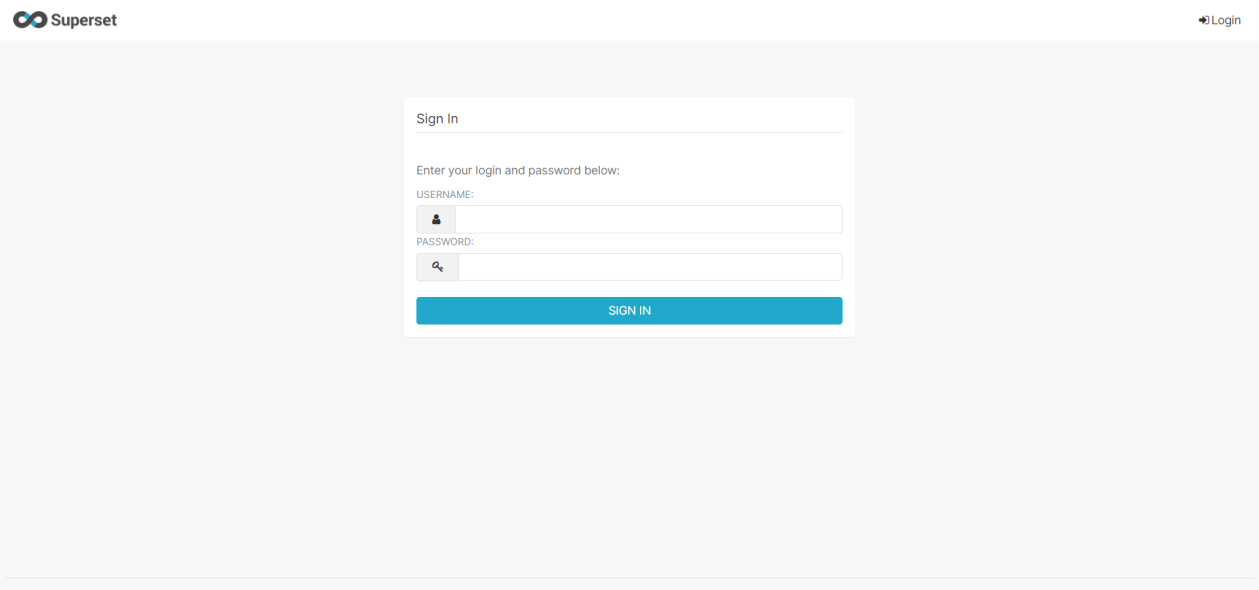
**--timeout：worker进程超时时间，超时会自动重启**

**--bind：绑定本机地址，即为Superset访问地址**

**--daemon：后台运行**

**（3）登录Superset**

访问[http://master:8787](http://hadoop102:8787)，并使用2.2.2节中第4步创建的管理员账号进行登录。



**3）停止superset**

**（1）停掉**gunicorn进程

(superset) [test1@master ~]$ ps -ef | awk '/superset/ && !/awk/{print $2}' | xargs kill -9

**（2）退出superset环境**

(superset) [test1@master ~]$ conda deactivate

### 2.2.4 superset启停脚本

**1）创建superset.sh文件**

[test1@master bin]$ vim superset.sh

**内容如下**

#!/bin/bash

superset\_status(){

result=`ps -ef | awk '/gunicorn/ && !/awk/{print $2}' | wc -l`

if [[ $result -eq 0 ]]; then

return 0

else

return 1

fi

}

superset\_start(){

source ~/.bashrc

superset\_status >/dev/null 2>&1

if [[ $? -eq 0 ]]; then

conda activate superset ; gunicorn --workers 5 --timeout 120 --bind master:8787 --daemon 'superset.app:create\_app()'

else

echo "superset正在运行"

fi

}

superset\_stop(){

superset\_status >/dev/null 2>&1

if [[ $? -eq 0 ]]; then

echo "superset未在运行"

else

ps -ef | awk '/gunicorn/ && !/awk/{print $2}' | xargs kill -9

fi

}

case $1 in

start )

echo "启动Superset"

superset\_start

;;

stop )

echo "停止Superset"

superset\_stop

;;

restart )

echo "重启Superset"

superset\_stop

superset\_start

;;

status )

superset\_status >/dev/null 2>&1

if [[ $? -eq 0 ]]; then

echo "superset未在运行"

else

echo "superset正在运行"

fi

esac

**2）加执行权限**

[test1@master bin]$ chmod +x superset.sh

**3）测试**

**（1）启动superset**

[test1@master bin]$ superset.sh start

**（2）停止superset**

[test1@master bin]$ superset.sh stop

# 第3章 Superset使用

## 3.1 对接MySQL数据源

### 3.1.1 安装依赖（前边步骤已安装，可跳过）

(superset) [test1@master ~]$ conda install mysqlclient

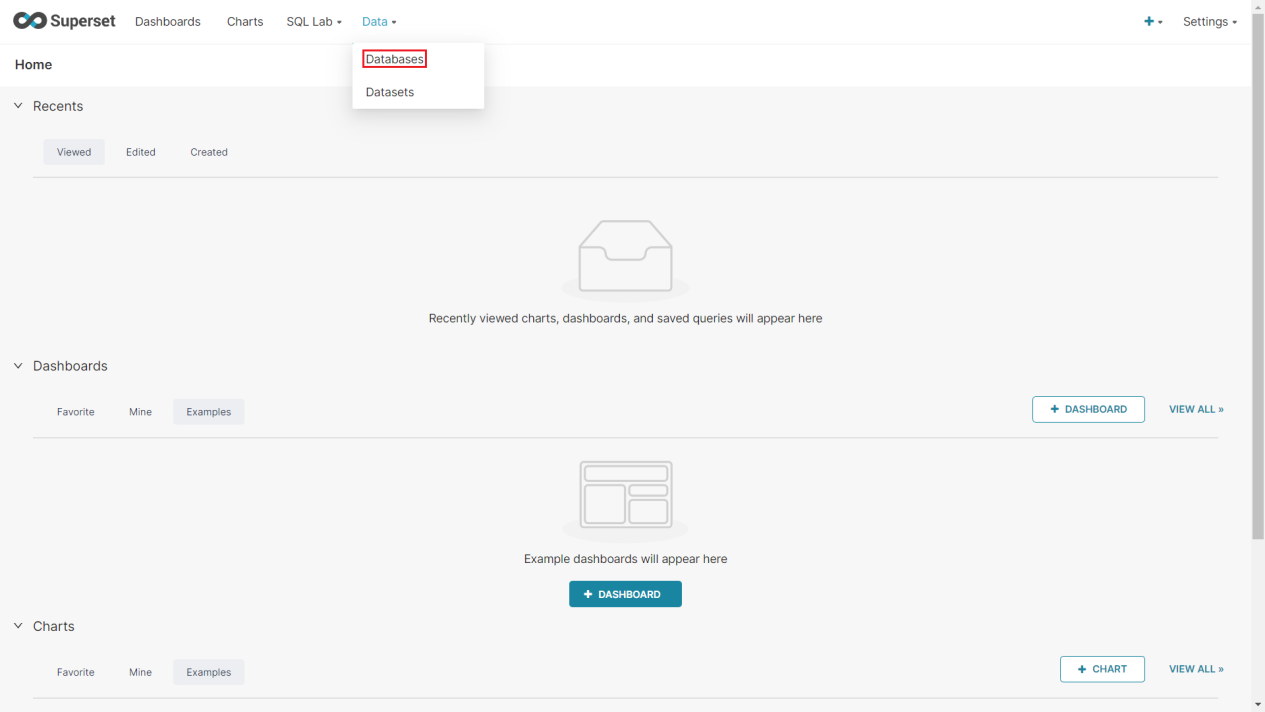
**说明：对接不同的数据源，需安装不同的依赖，以下地址为官网说明。**

**<https://superset.apache.org/docs/databases/installing-database-drivers>**

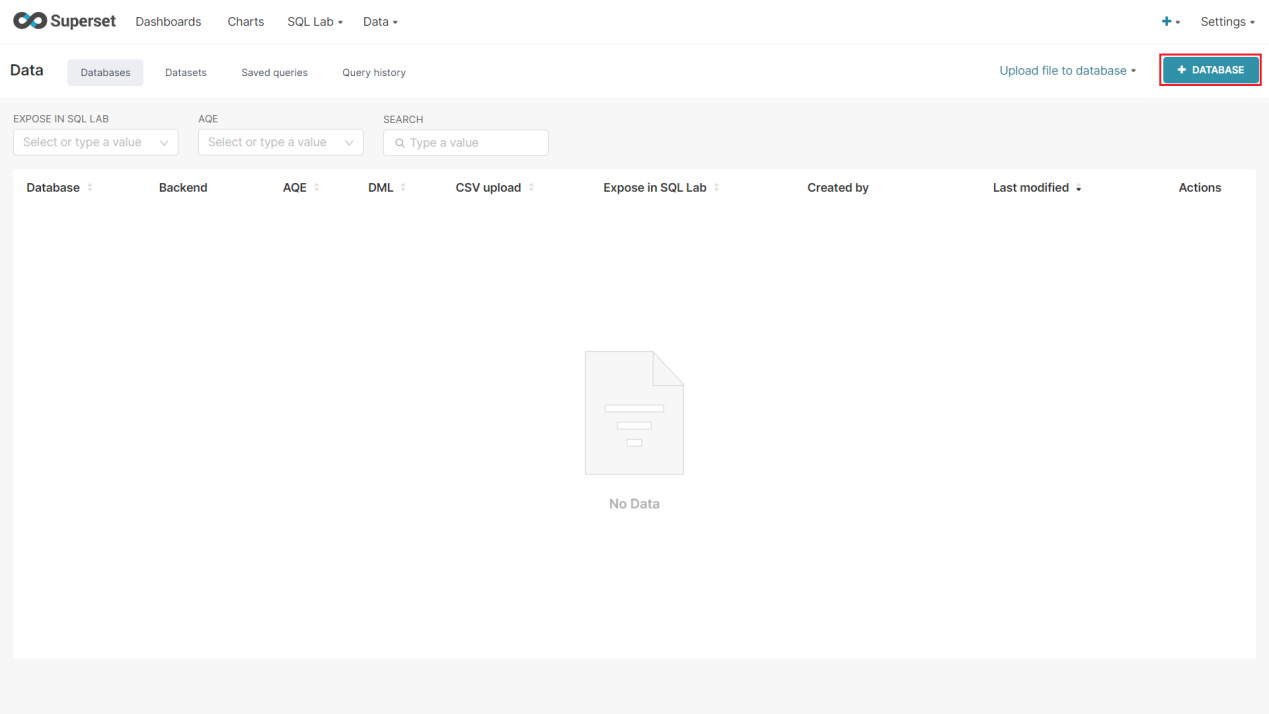
### 3.1.2 数据源配置

**1）Database配置**

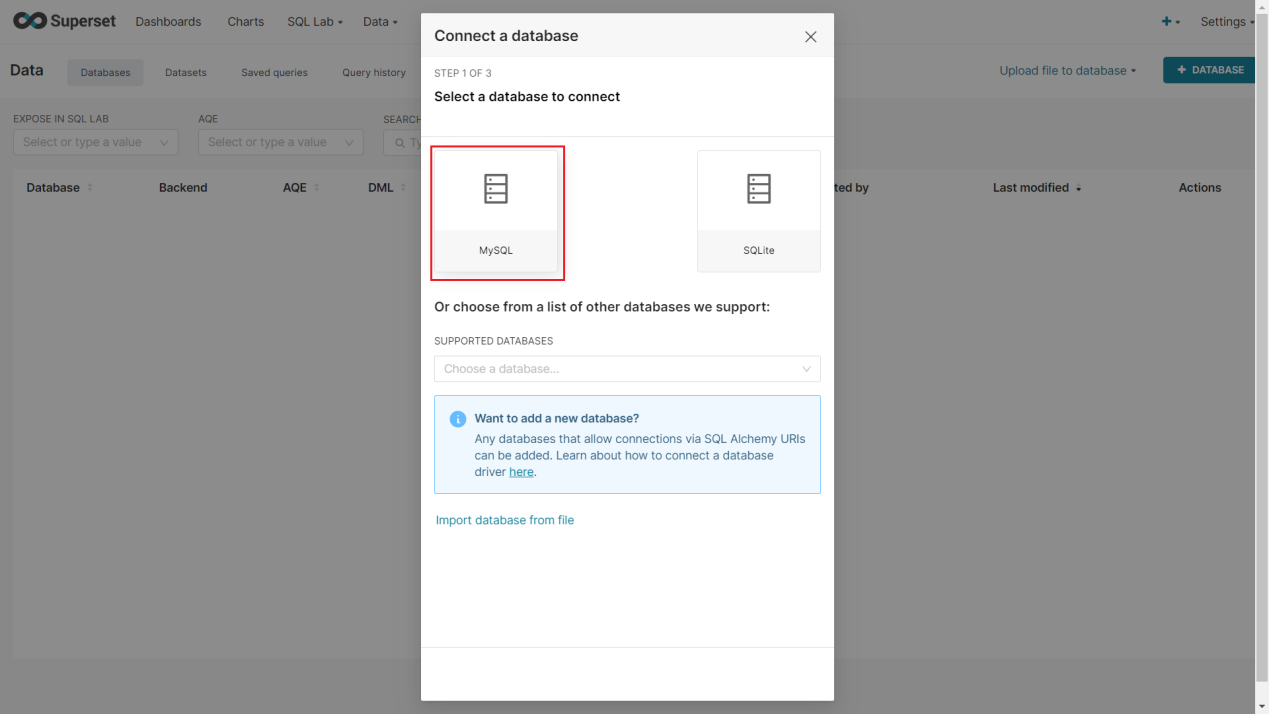
**Step1：**点击Data/Databases



**Step2：**点击**＋DATABASE**



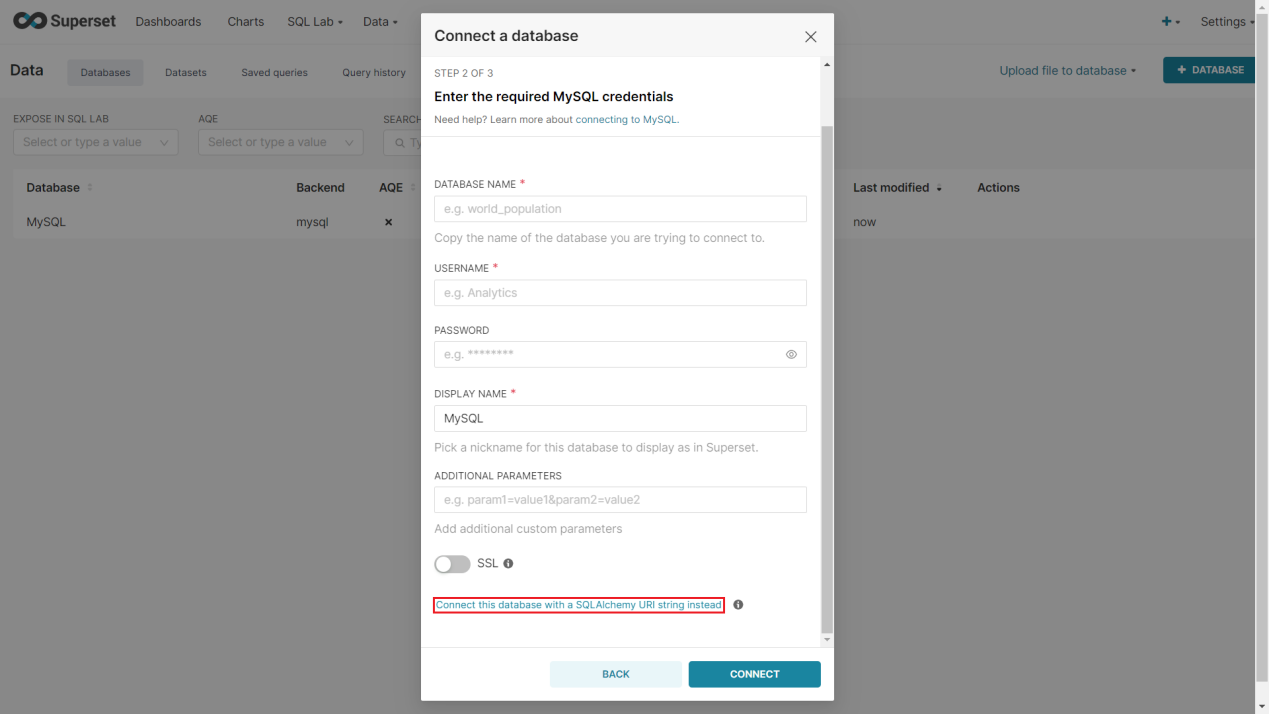
**Step3：**点击填写Database及SQL Alchemy URI

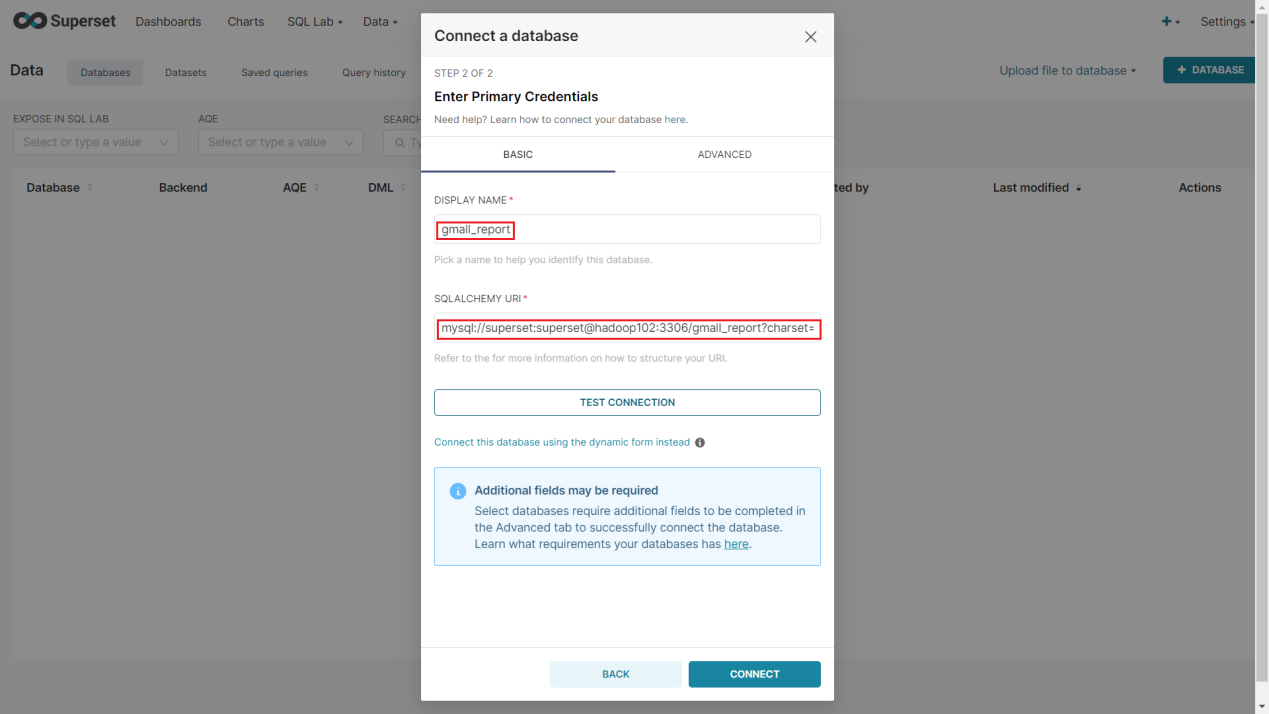


（1）方式一：通过URL连接

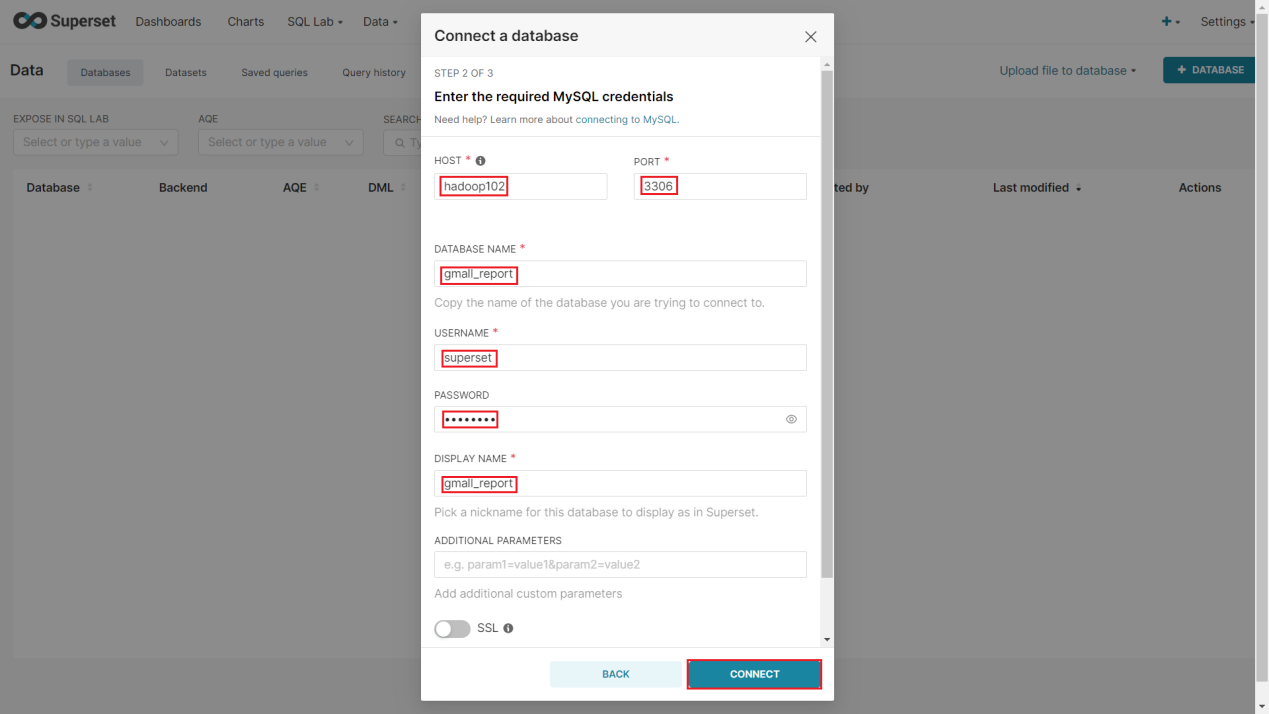
注：SQL Alchemy URI编写规范：mysql://用户名:密码@主机名:端口号/数据库名称

此处填写：mysql://superset:superset@master:3306/gmall\_report?charset=utf8



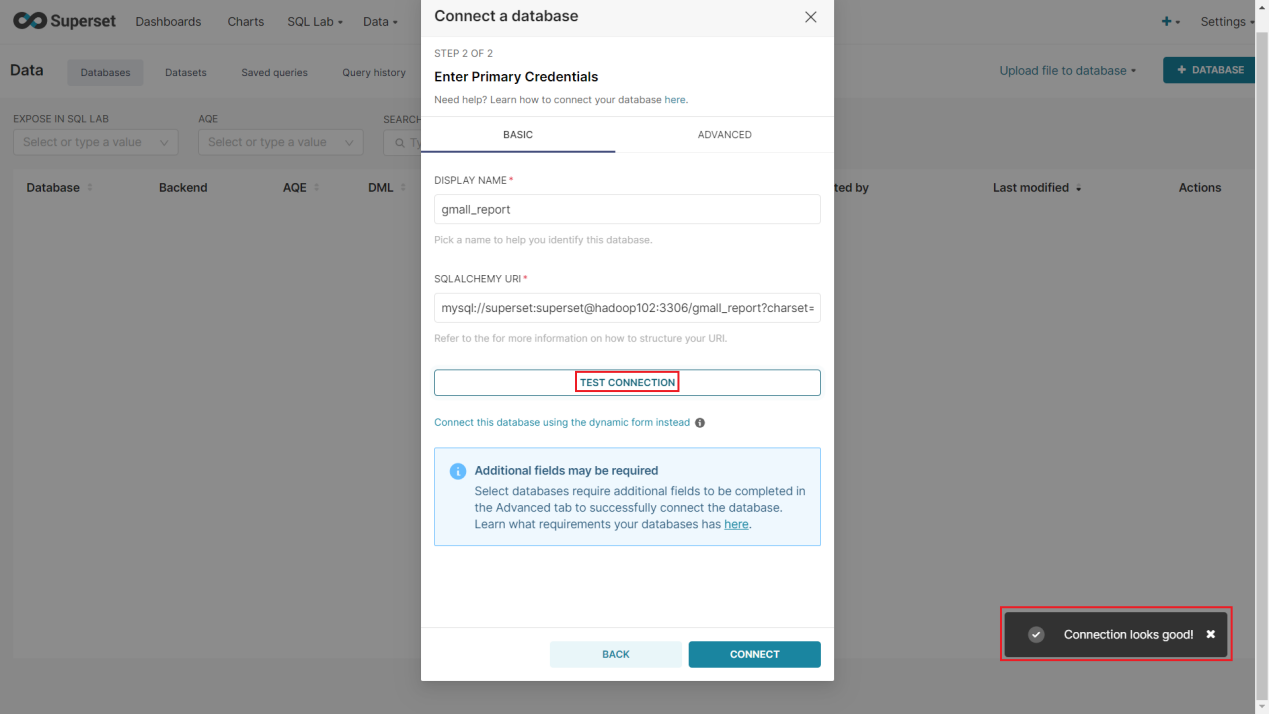


（2）方式二：逐个输入认证信息

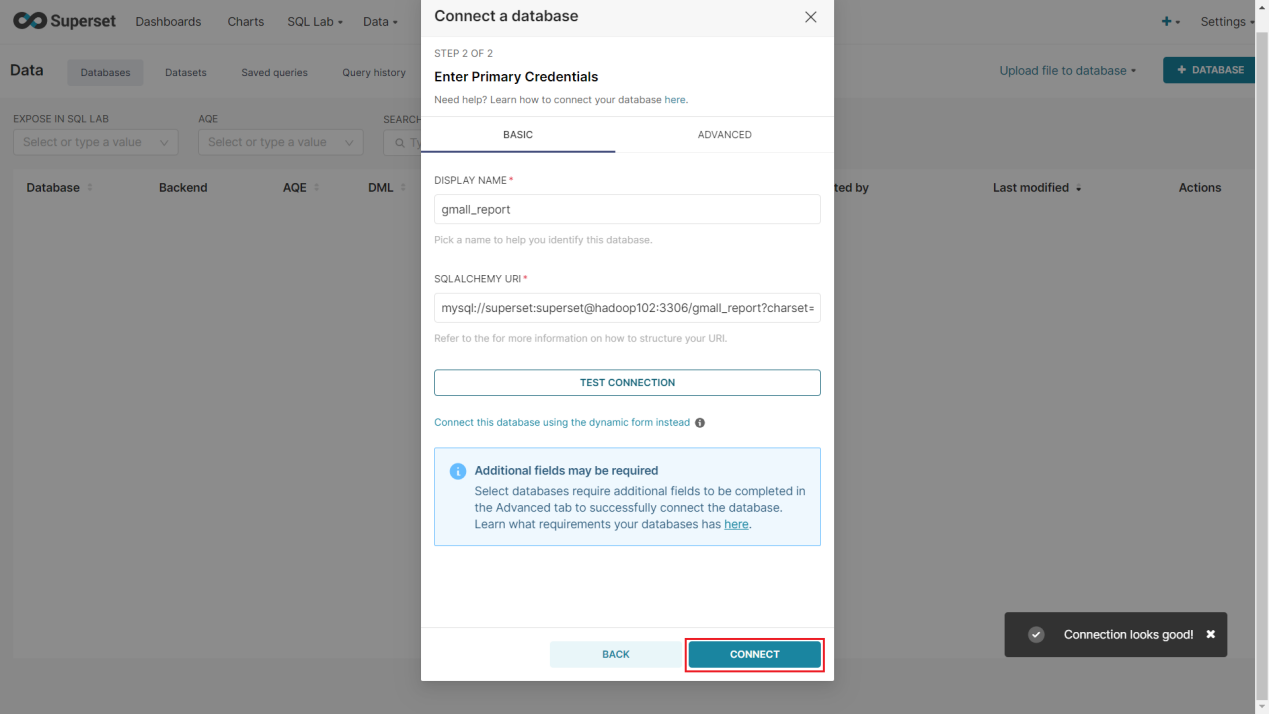


此处选择方式一。

**Step4：**点击Test Connection，出现“Connection looks good!”提示即表示连接成功。

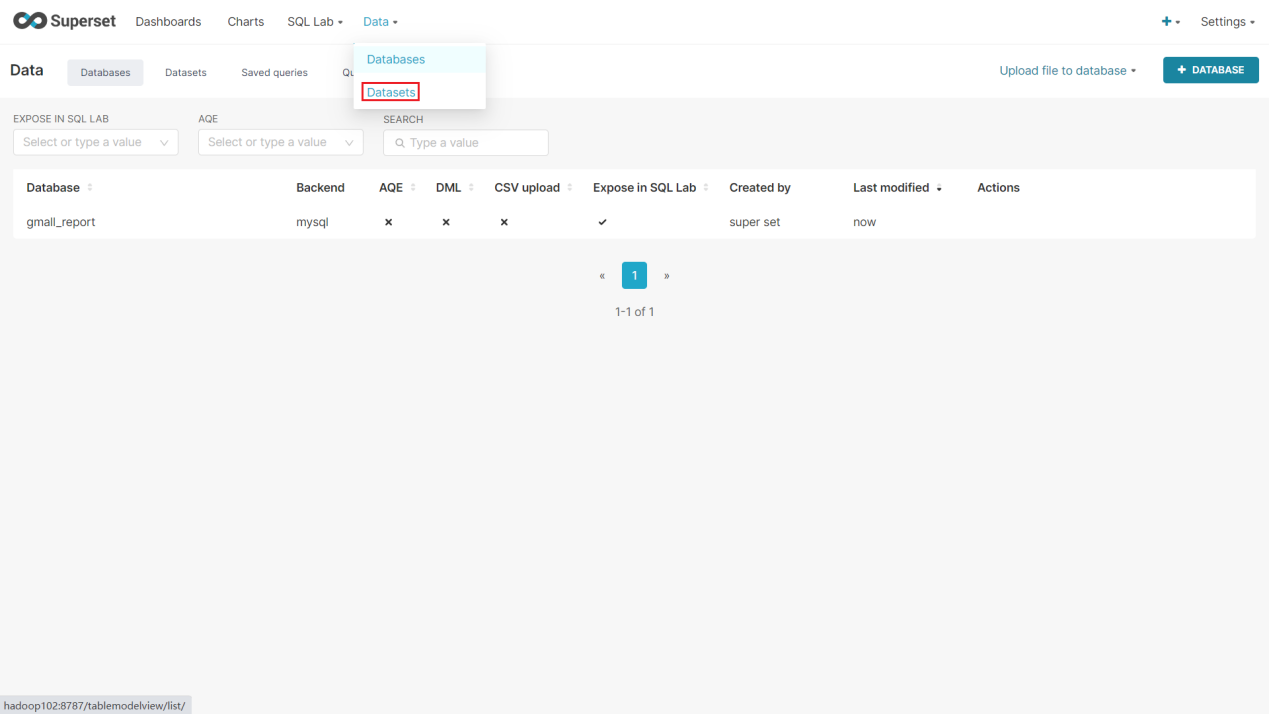


**Step5：**点击CONNECT

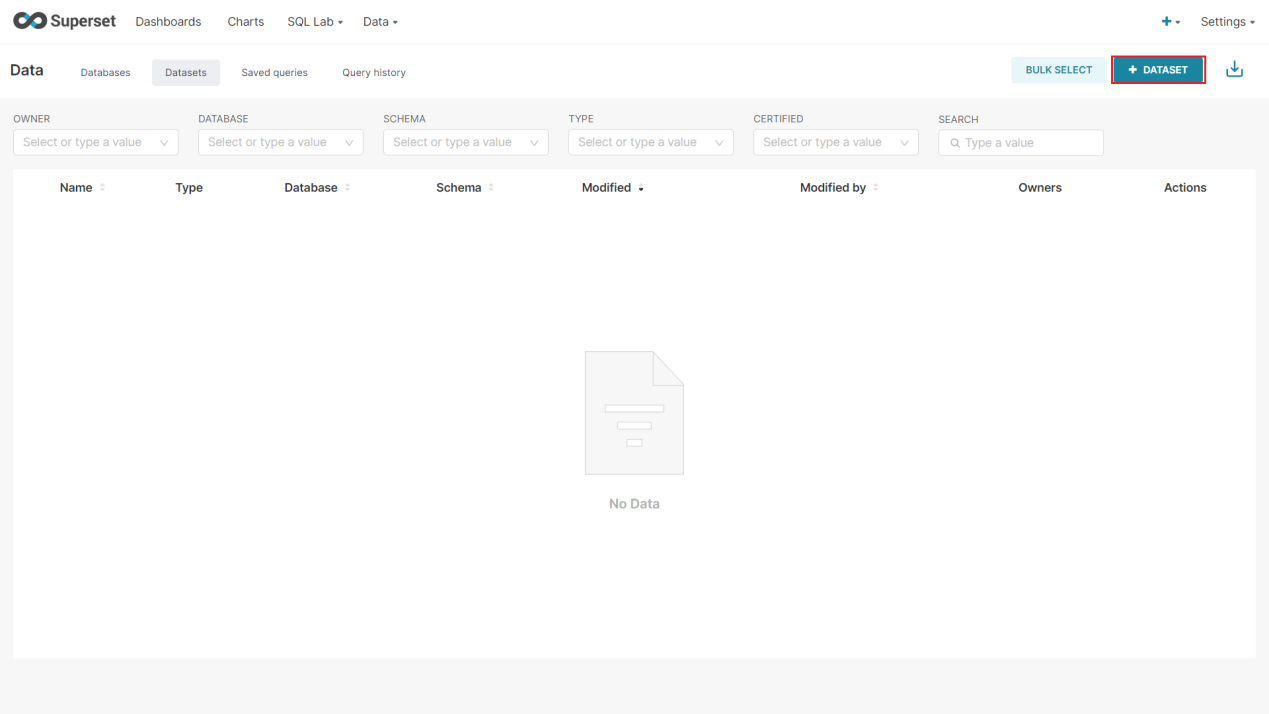


**2）Table配置**

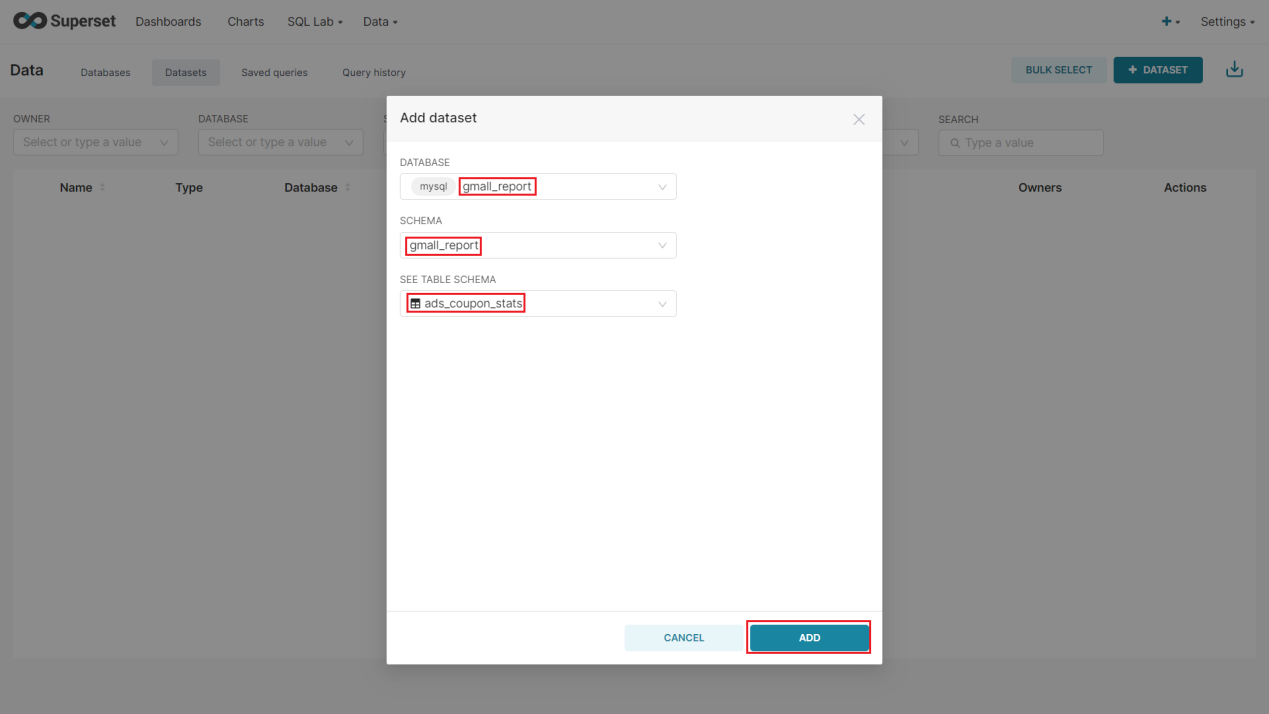
**Step1：点击**Data/Datasets



**Step2：点击+DATASET**



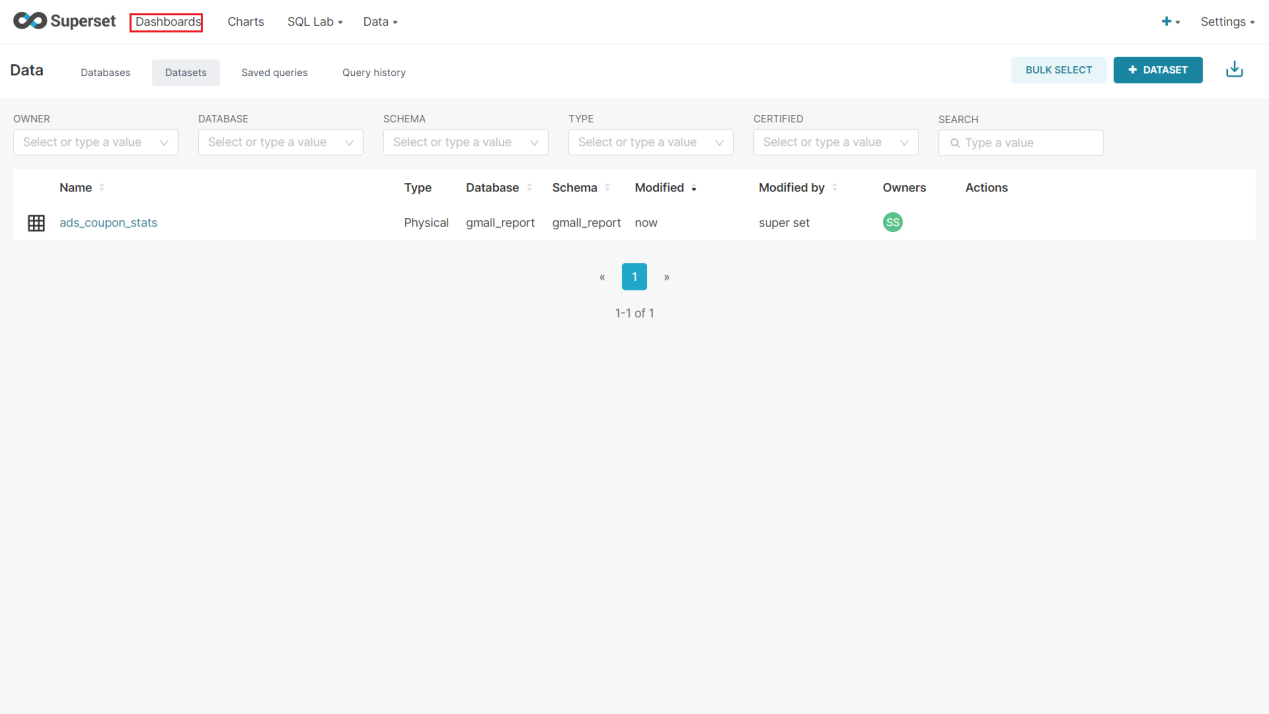
**Step3：配置Table**

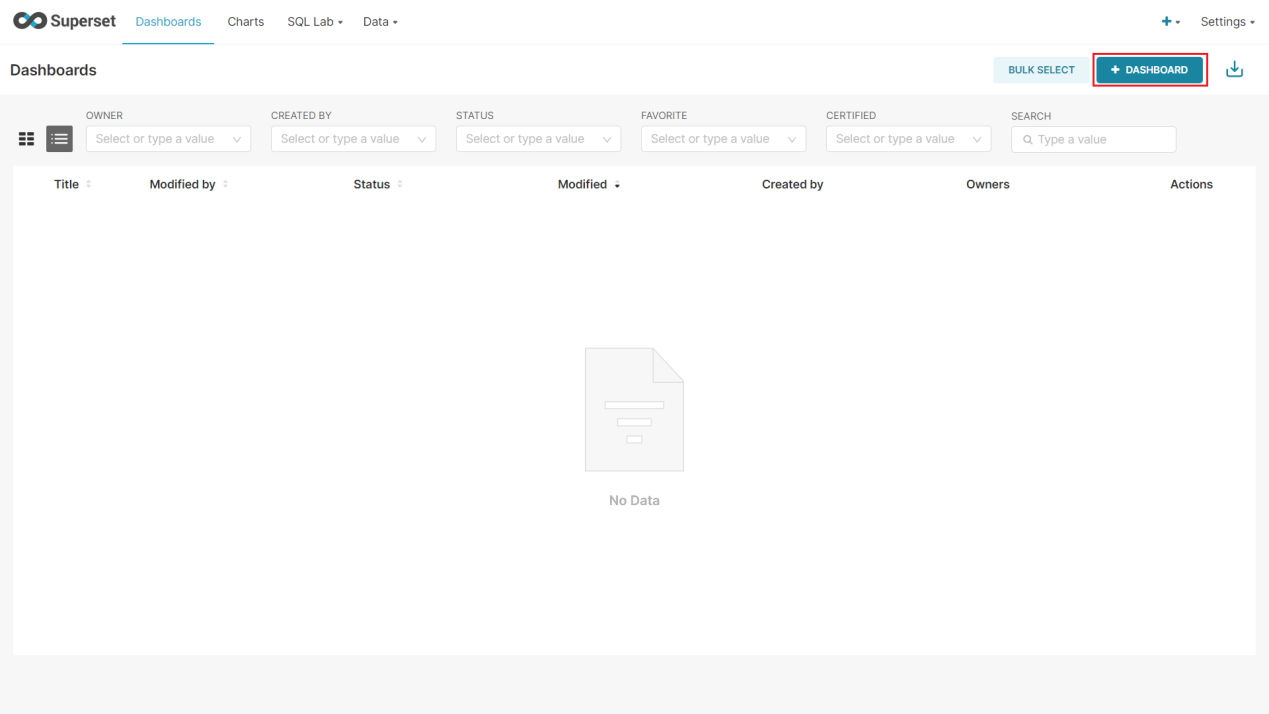


## 3.2 制作仪表盘

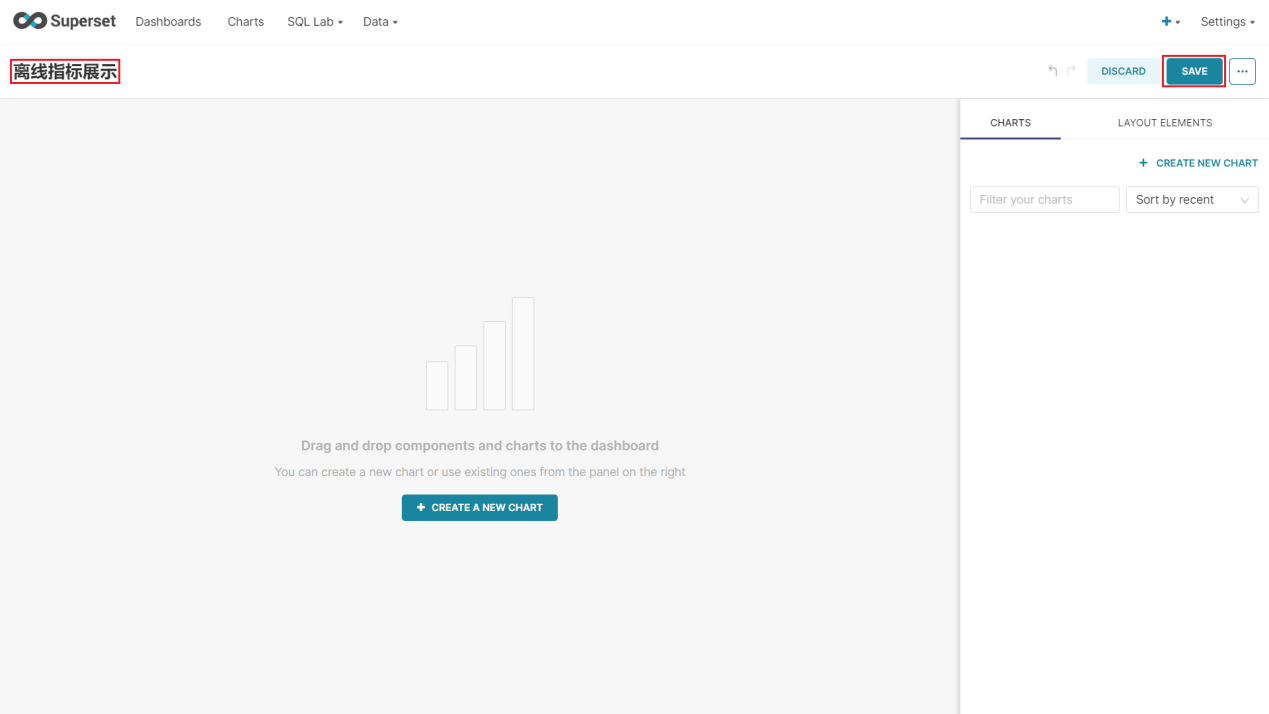
### 3.2.1 创建空白仪表盘

**1）点击Dashboards/+DASHBOARDS**



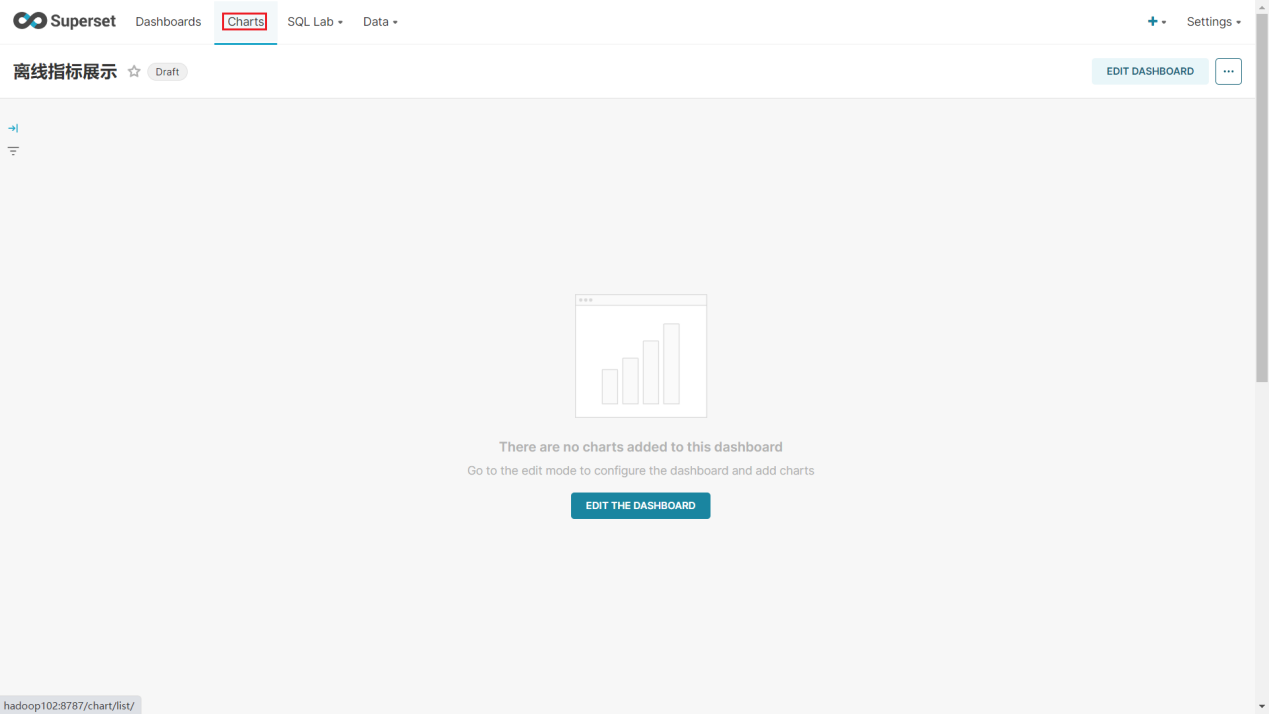


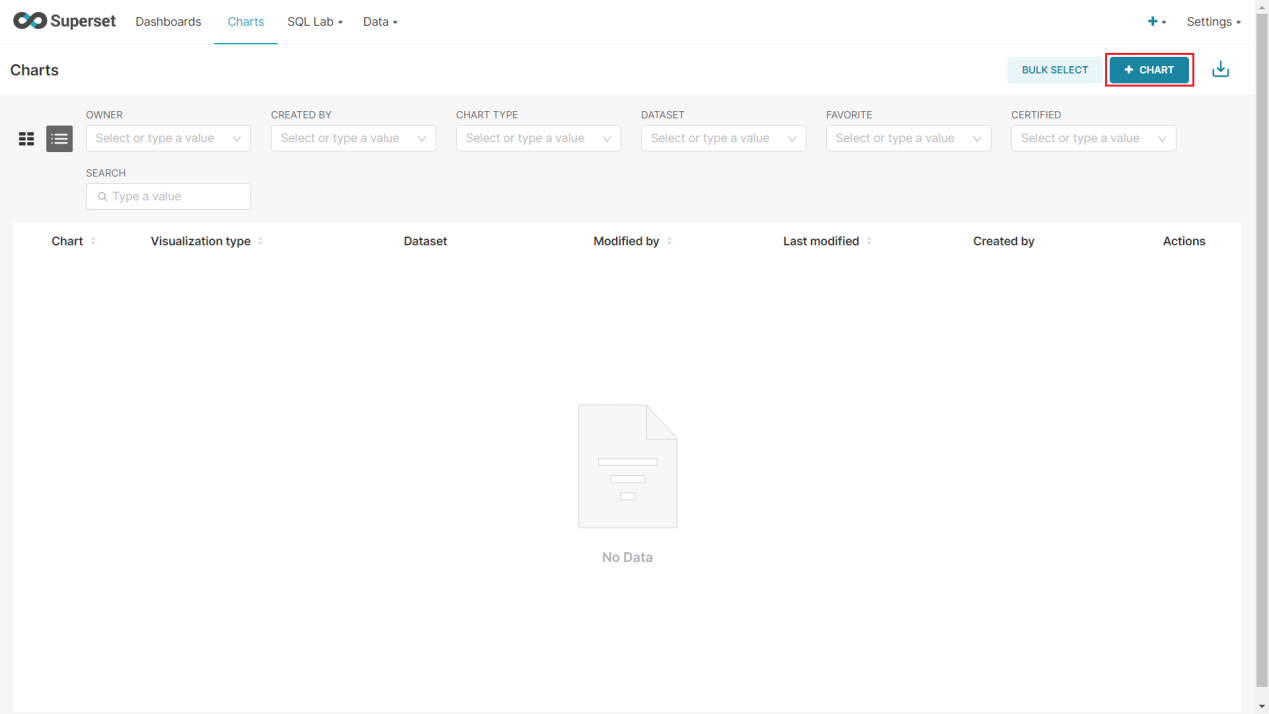
**2）命名并保存**



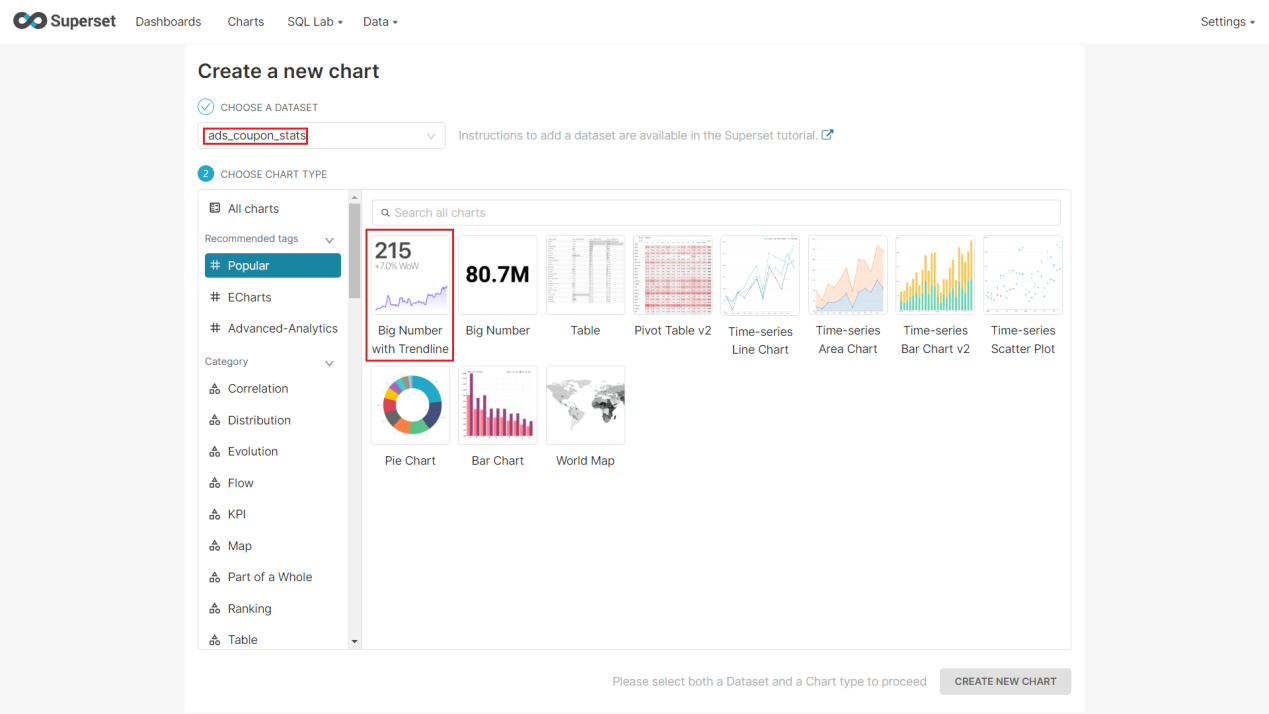
### 3.2.2 创建图表

**1）点击Charts/+CHART**

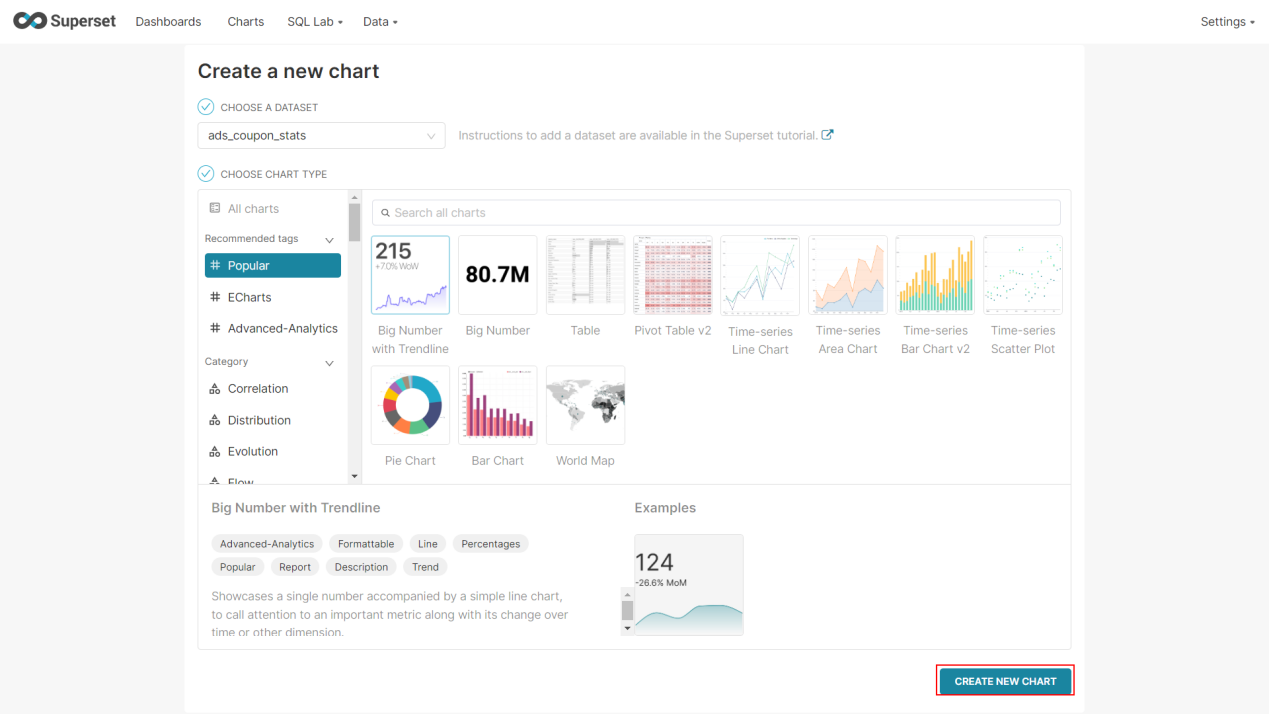




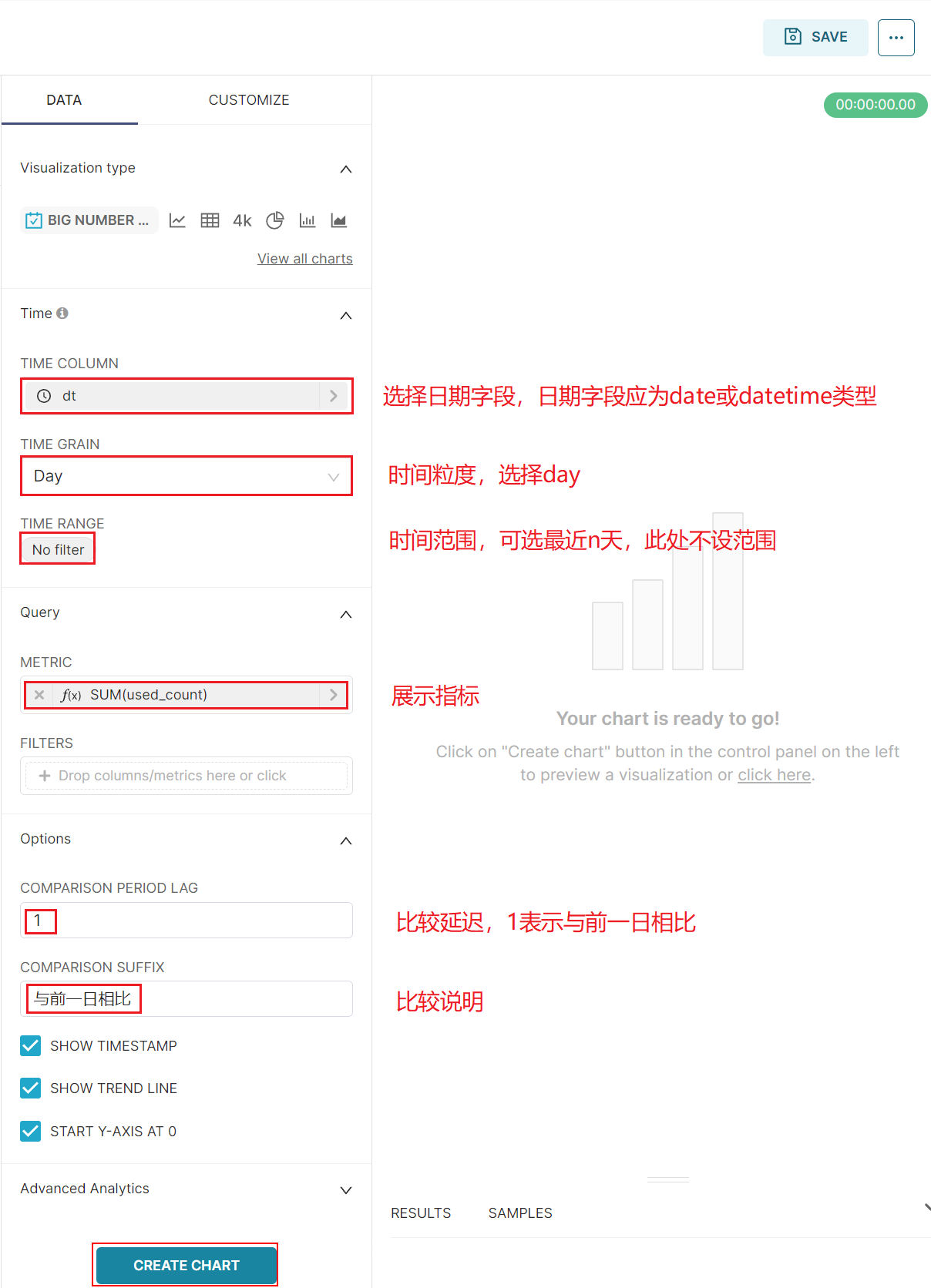
**2）选则数据源及图表类型**



**3）创建图表**



**4）按照说明配置图表并创建**

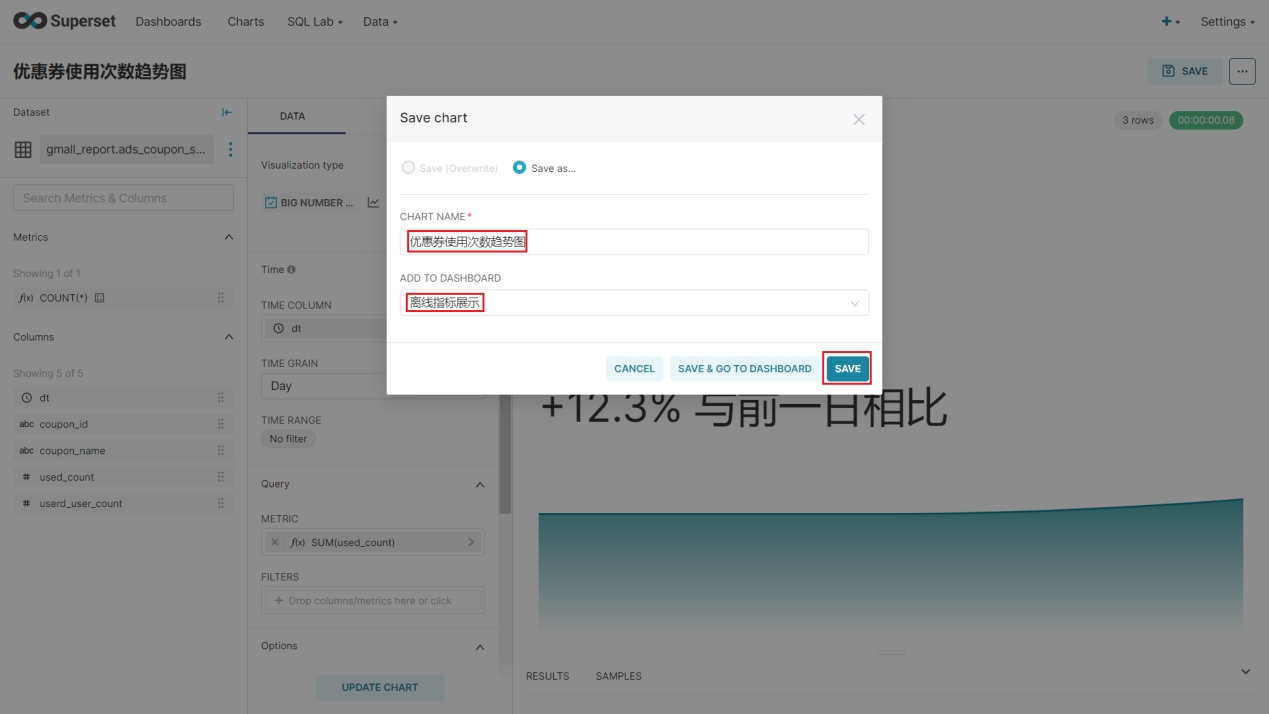


**5）如配置无误，可出现以下图标**



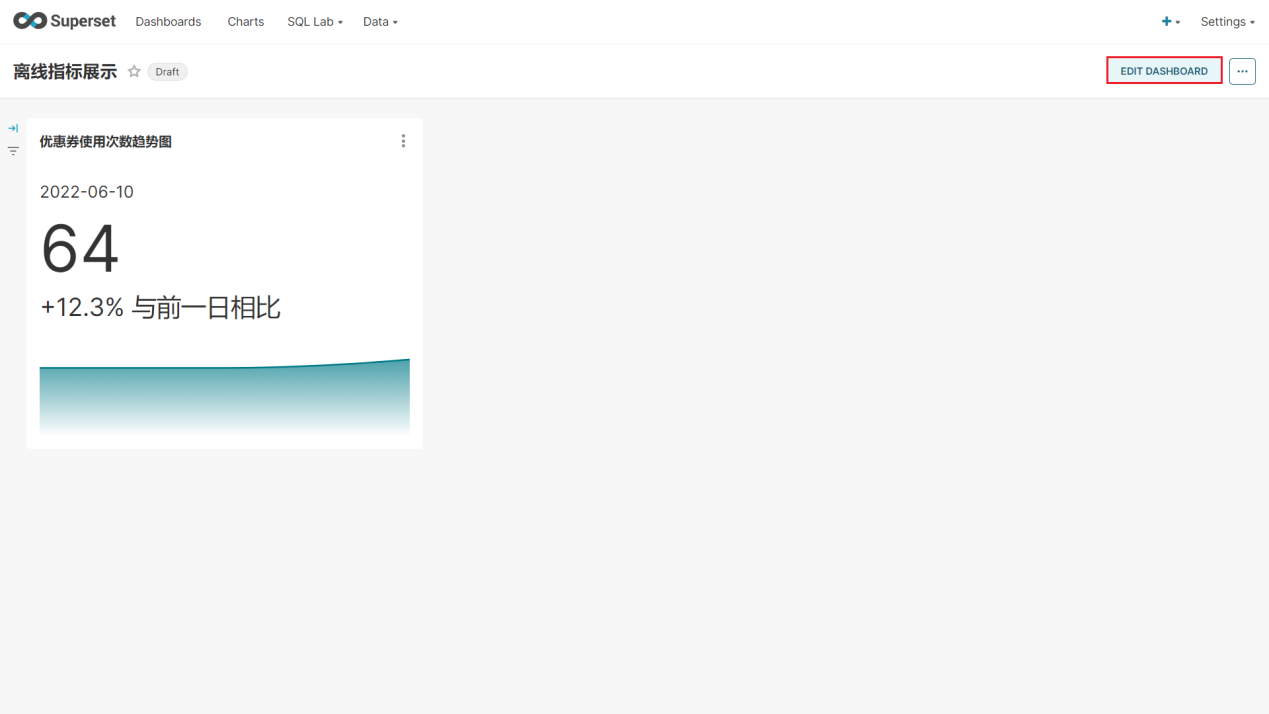
**8）命名该图表，并保存至仪表盘**



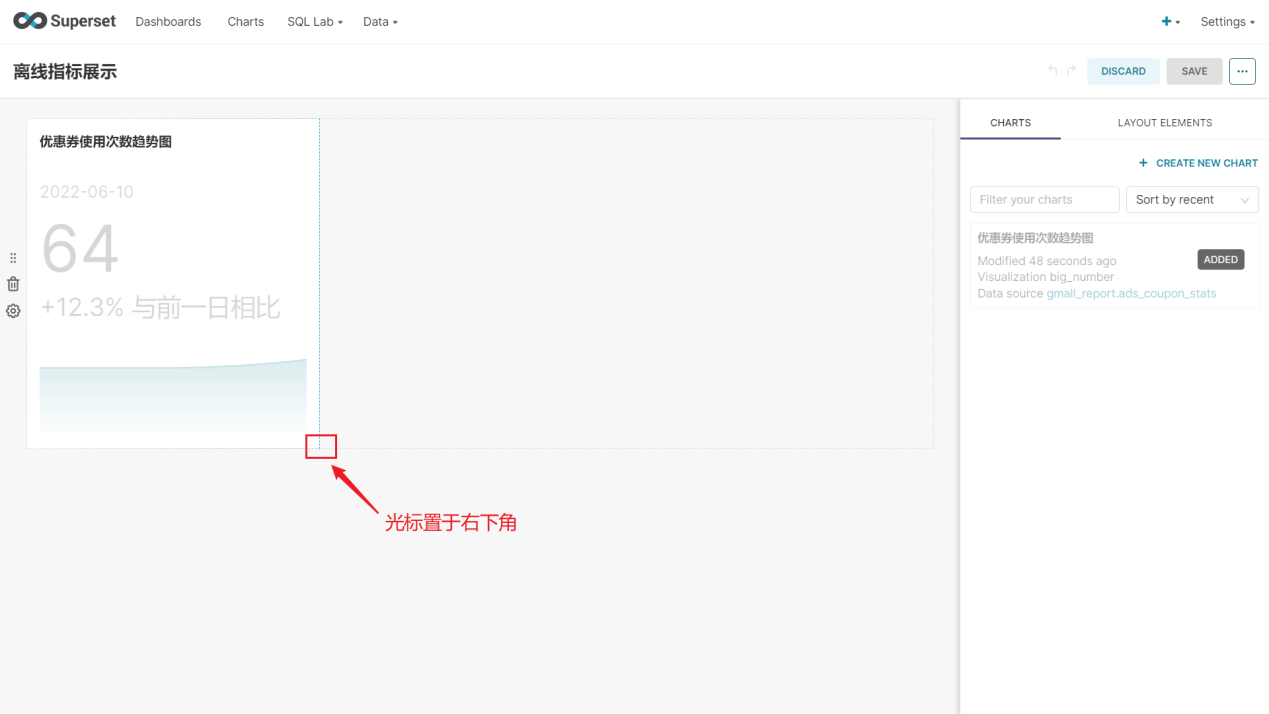


### 3.2.3 编辑仪表盘

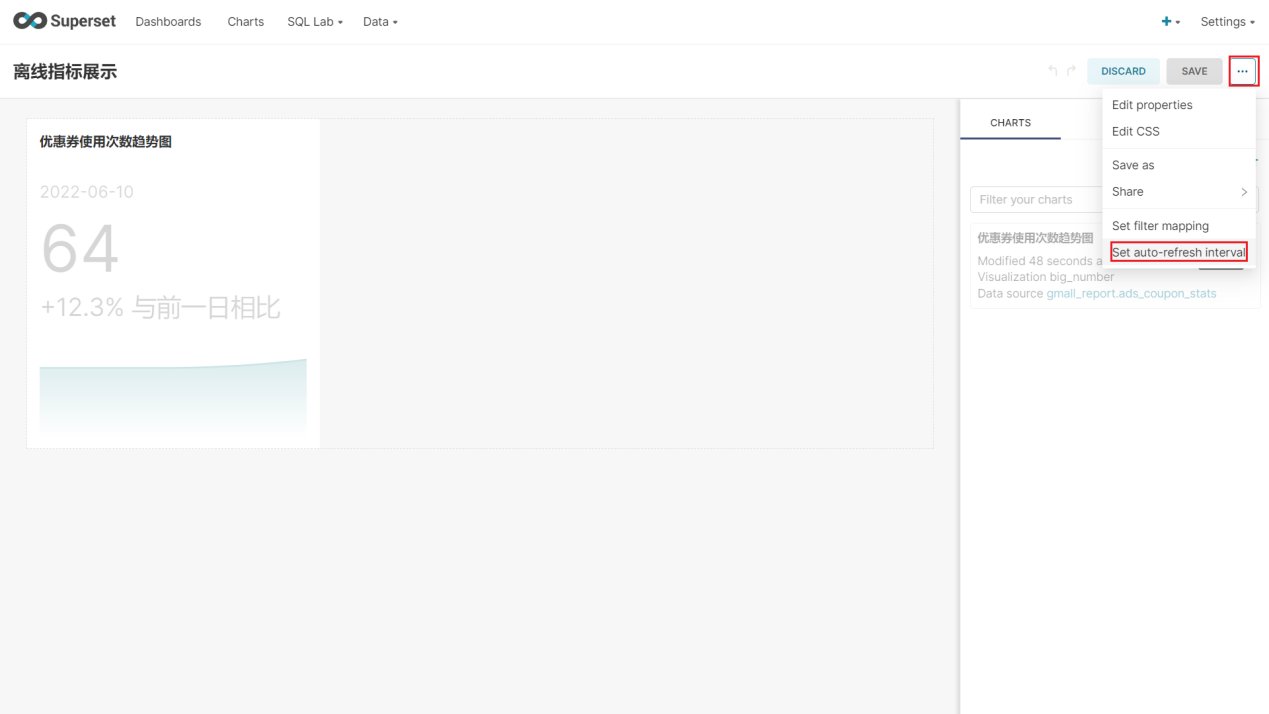
**1）打开仪表盘，点击编辑按钮**

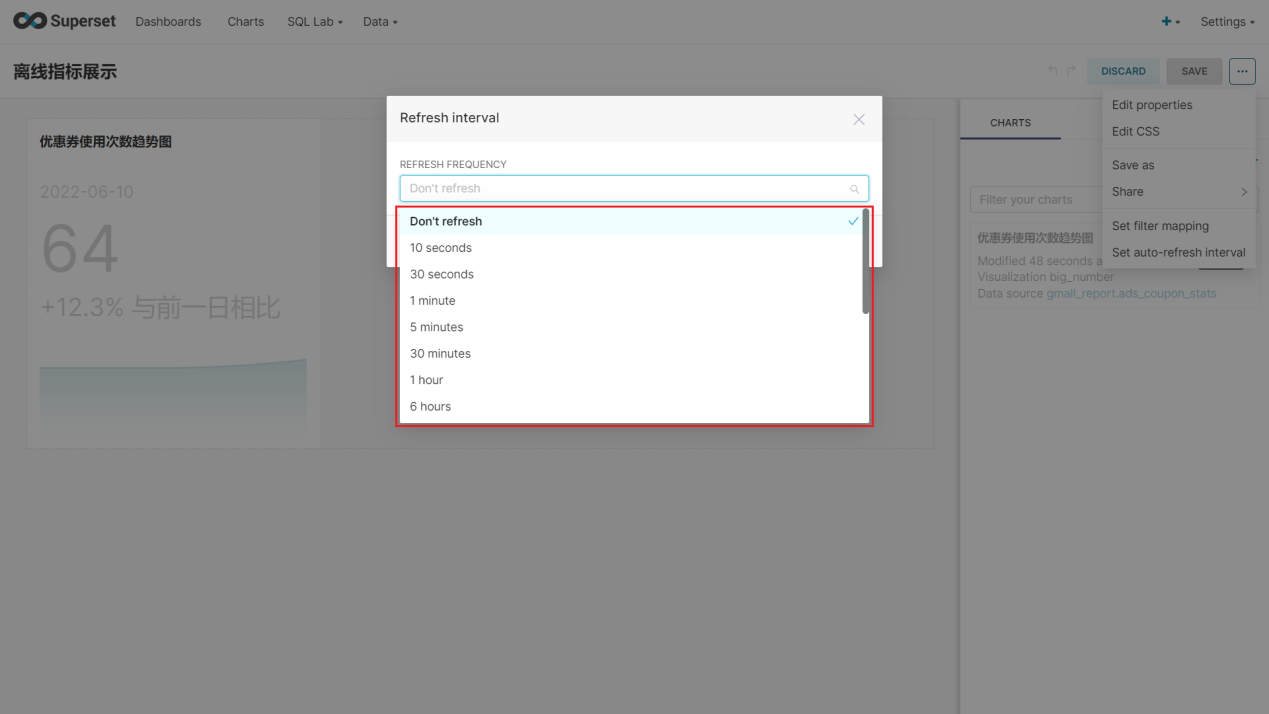


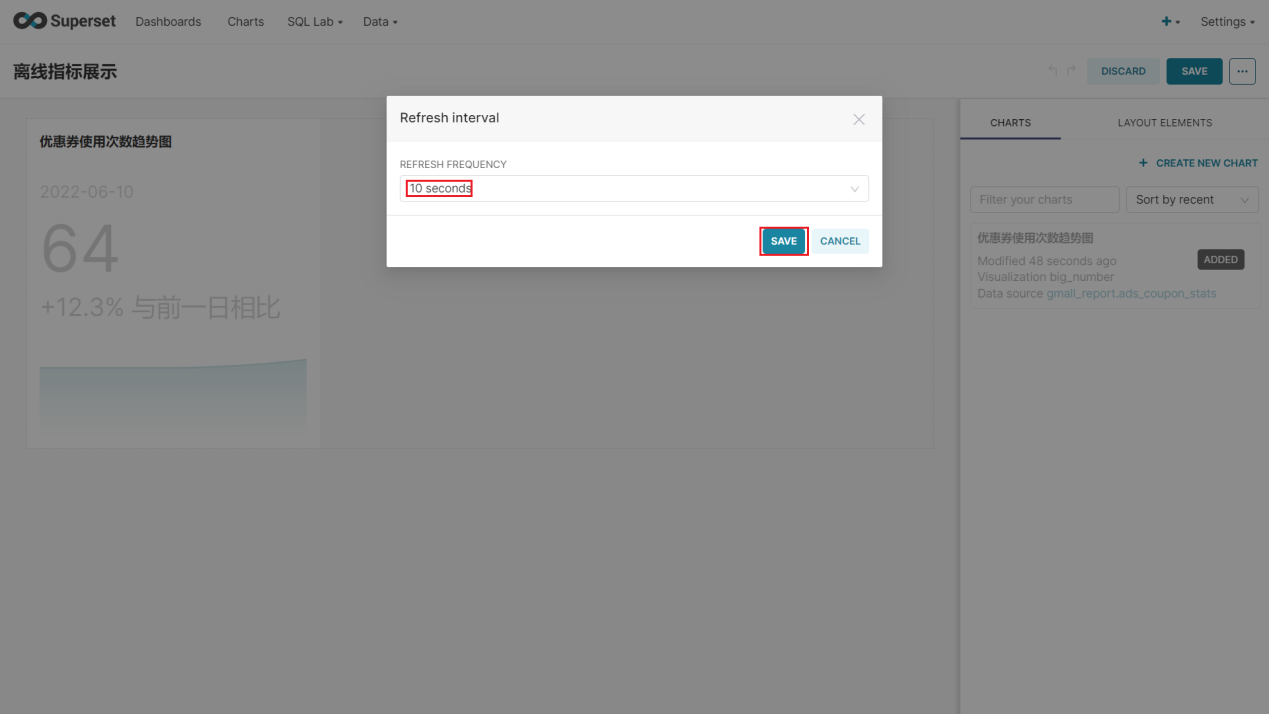
**2）调整图表大小以及图表盘布局**



**3）点击下图中箭头，可调整仪表盘自动刷新时间**



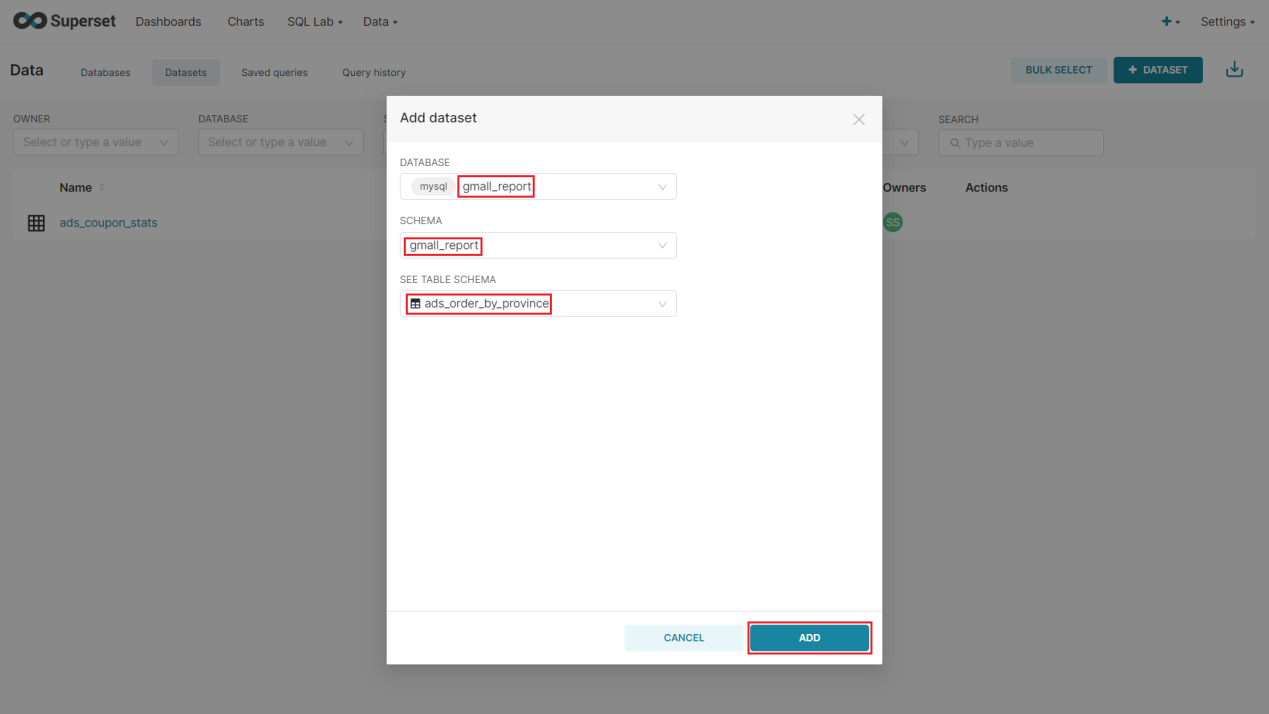




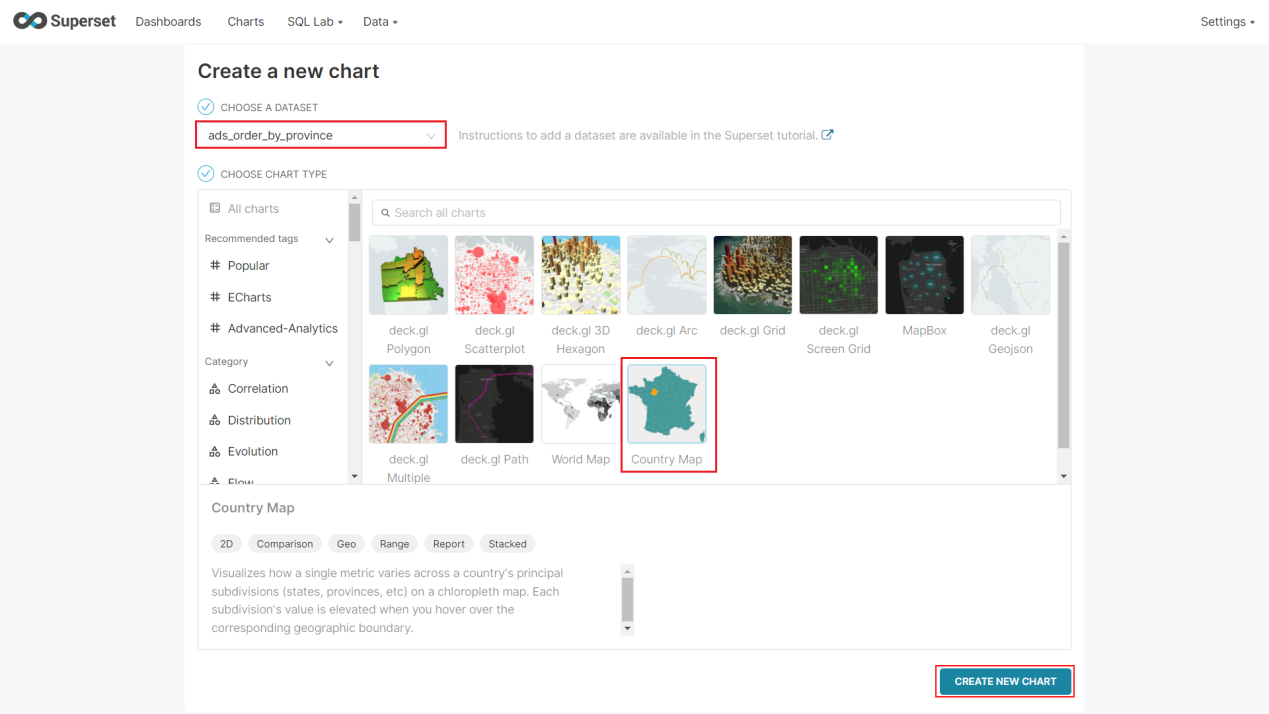
# 第4章 Superset实战

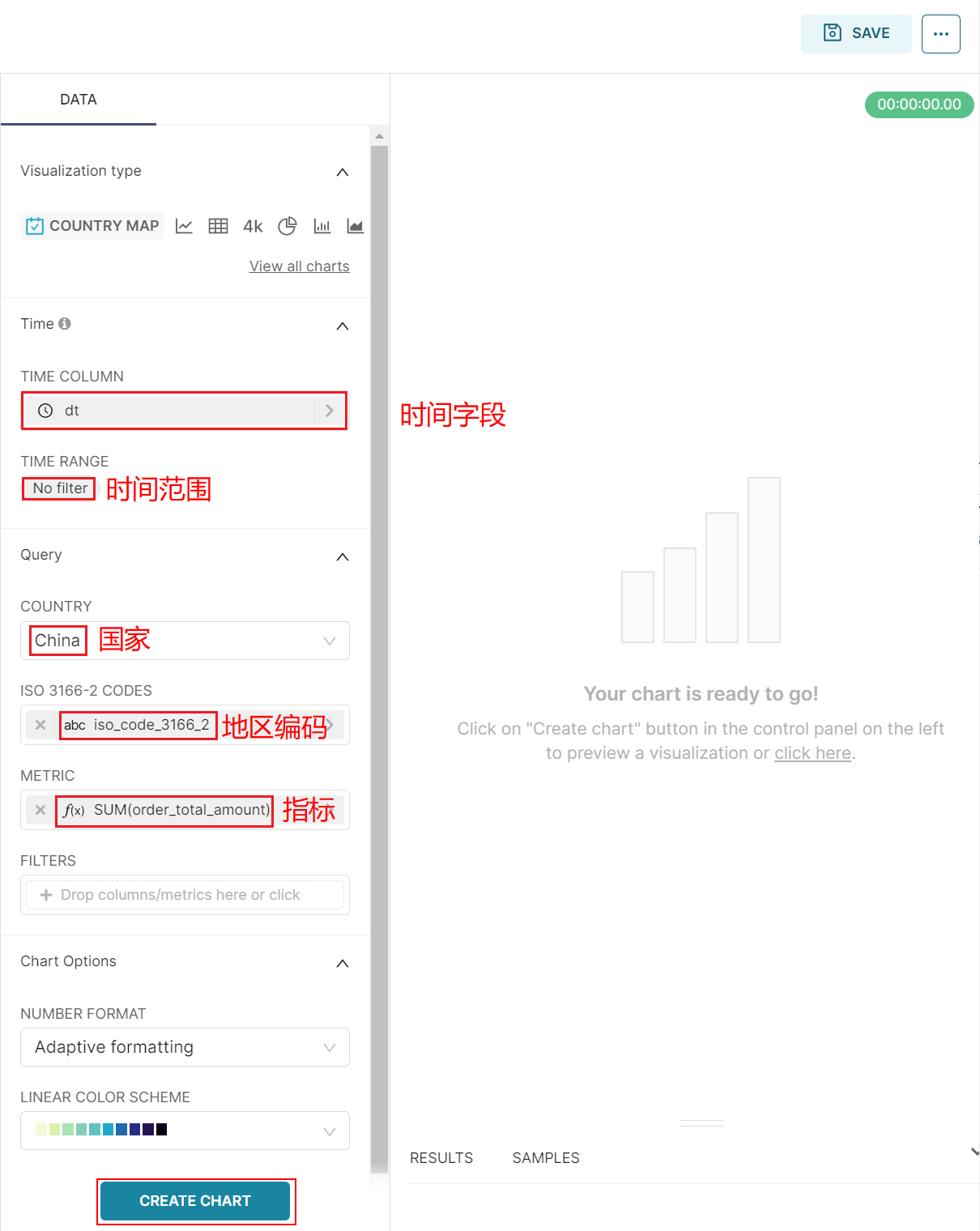
## 4.1 制作地图

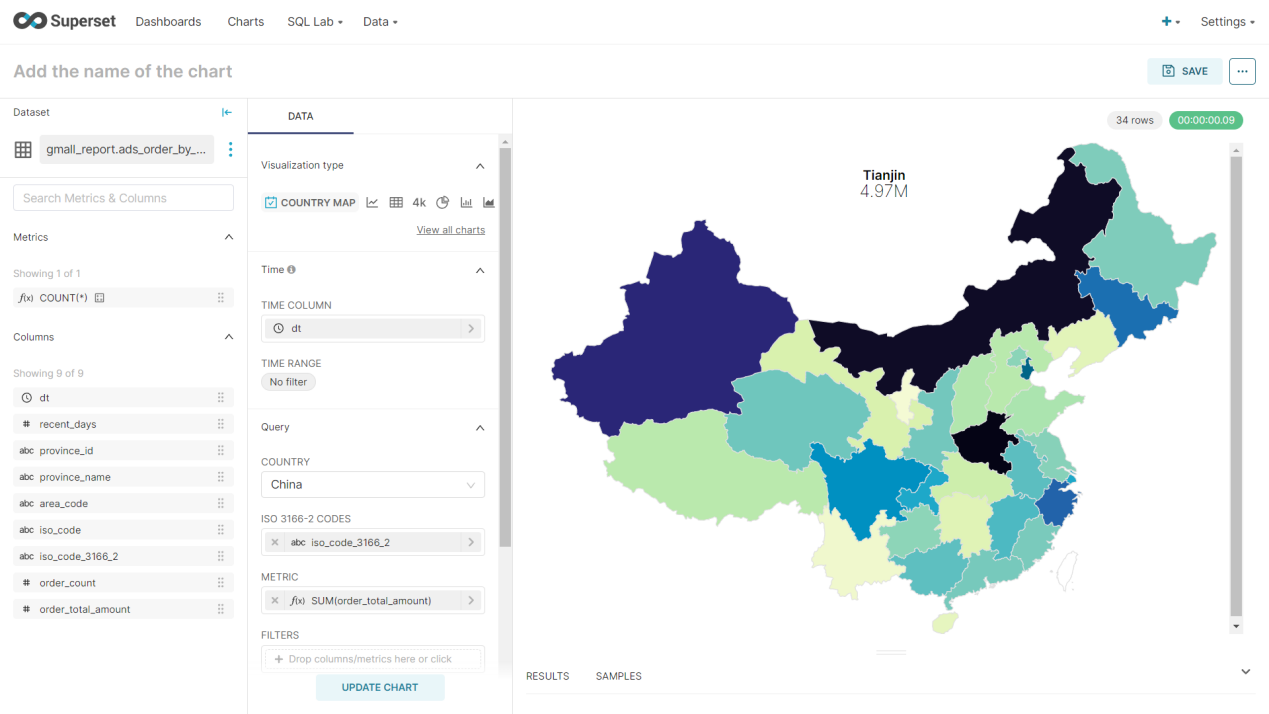
**4.1.1 配置Table**



**4.1.2 配置Chart**

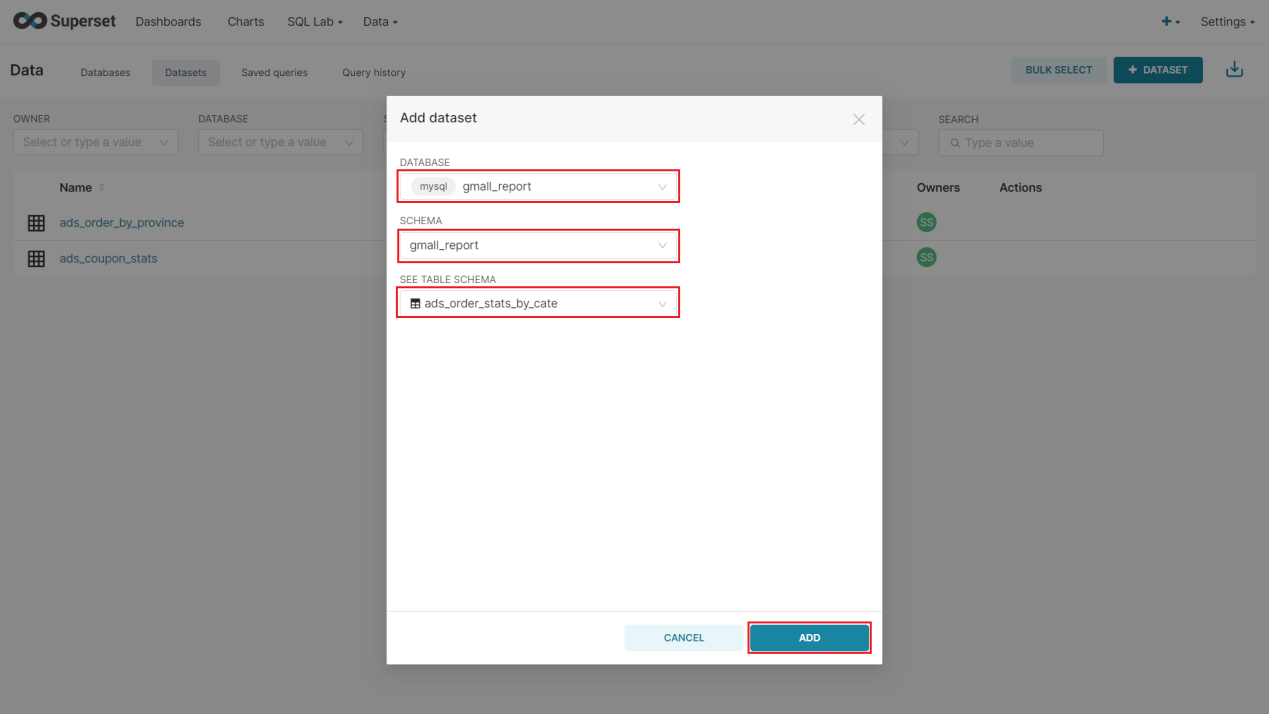






## 4.2 制作饼状图

**4.2.1 配置Table**



**4.2.2 配置Chart**

