Supersets Install 및 차트 생성

(이삭엔지니어링 인턴 김형근 fnfn9087@gmail.com)  
(10.100.3.246 VM에서 진행, Hive테이블은 DIA 클러스터의 것을 사용)

제공 환경 (VM) :   
Core : **16 Processor**  
RAM : **60GB**

설치 위치 : 10.100.3.246 (VM) /opt 폴더

**목차**

1. **설치 과정**
2. **Supersets 차트 생성  
   1) 데이터 가져오기**

**2) 차트 생성하기**

1. **설치 과정**
2. **다음과 같이 입력한다.**

sudo yum upgrade python-setuptools

sudo yum install gcc gcc-c++ libffi-devel python-devel python-pip python-wheel openssl-devel libsasl2-devel openldap-devel

1. **Python 3.6.x 설치**  
    sudo yum install <https://centos7.iuscommunity.org/ius-release.rpm>

sudo yum install python36u  
  
python3.6 -V //python 3.6.x 와 같이 상세버전이 표시된다.

sudo yum install python36u-pip  
sudo yum install python36u-devel

1. **Python virtualenv(가상환경)을 설정한다. -> superset은 python3.6이상 지원.  
   (supersets 설치가 끝날때까지 가상환경을 유지할 것!)**

pip install virtualenv

python3.6 -m venv venv //python3.6 -m venv (생성할 디렉토리 명)  
 source PATH/venv/bin/activate  
  
위와 같이 가상환경을 작동시키면 프롬프트 앞에 “(디렉토리명)” 표시가 뜬다.  
 ex) (venv)[root@worker1~] $  
  
그리고, $python -V를 입력하면 python 3.6.7로 뜨는 것을 확인할 수 있다.  
본래환경에서는 $python3.6 -V 라고 입력해야 python3이상의 버전확인이 가능했다.  
*본래환경 : python 2.7.5  
가상환경 : python 3.6.7*   
  
가상환경을 종료시키고 싶을땐, **$deactivate** 라고 입력하면 된다.

**Python2가 설치되어 있지만 Python3 환경에서만 실행되는 프로그램도 있을 수 있다.  
(Python2에서의 요소들이 Python3에서 새로 변경되었거나 지원을 안 하기 때문이다. )  
그럴 경우에는 virtualenv로 가상환경을 설정하여 2버전과 3버전을 따로 쓸 수 있는 독립적인 환경을 구성해주는 것이다.**

1. **Python’s setup tools and pip**

pip install --upgrade setuptools pip

1. **Supersets installation**

*# Install superset*

pip install superset

*# Create an admin user (you will be prompted to set a username, first and last name before setting a password)*

fabmanager create-admin --app superset

*# Initialize the database*

superset db upgrade

*# Load some data to play with*

superset load\_examples

*# Create default roles and permissions*

superset init

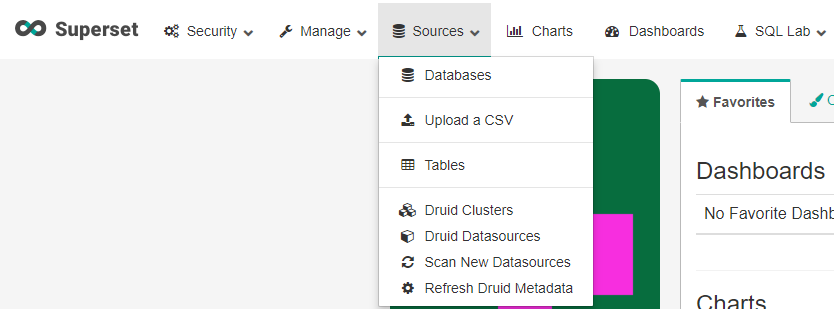
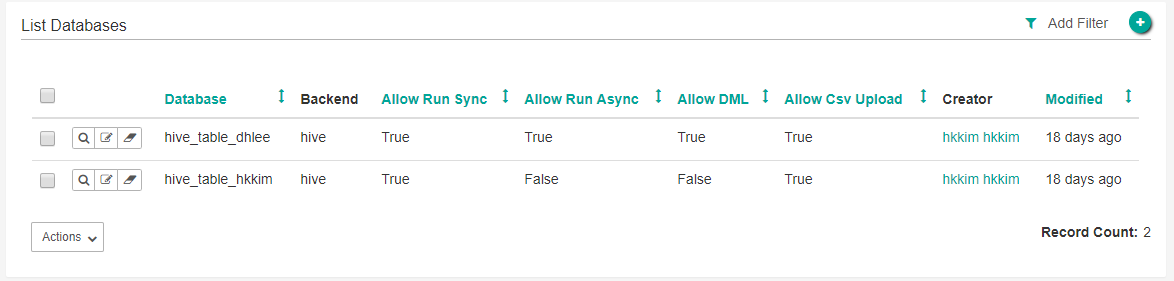
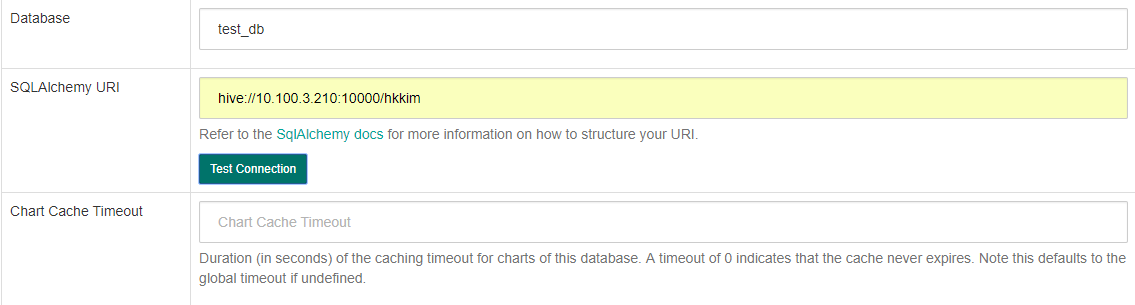
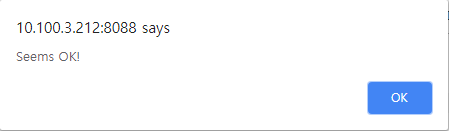
*# To start a development web server on port 8088, use -p to bind to another port*

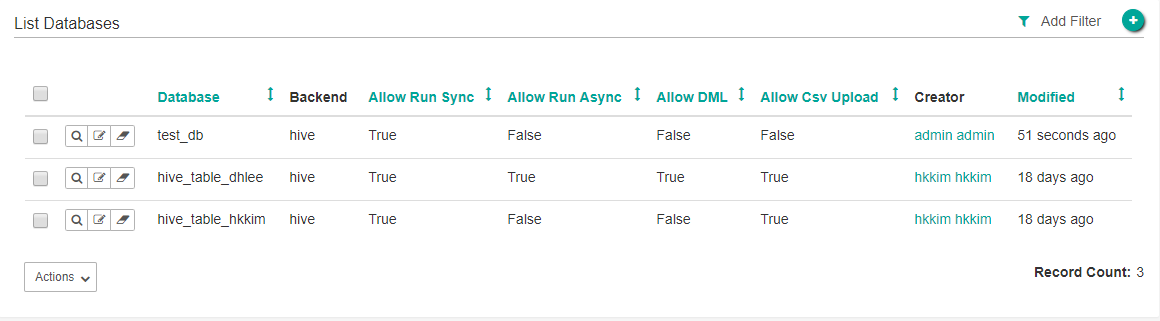
superset runserver

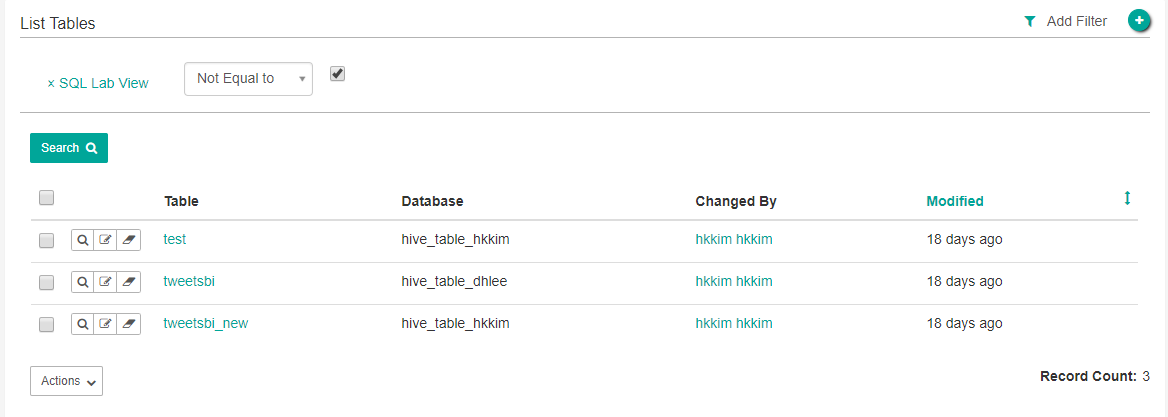
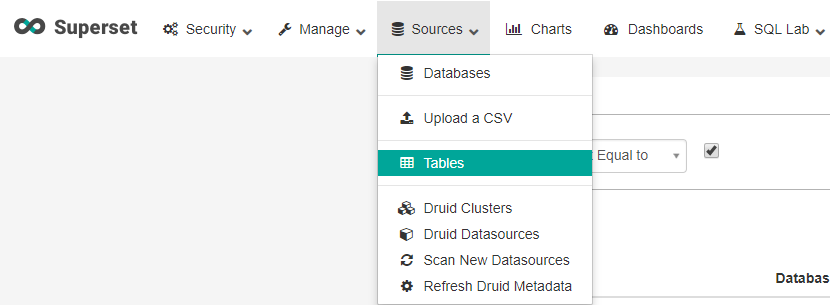
1. **SuperSets 접속(** [**http://10.100.3.246:8088**](http://10.100.3.246:8088) **) 후, Hive connect error 해결**  
     
   “Error in sasl\_client\_start (-4) SASL(-4): no mechanism available: No worthy mechs found”  
   라는 에러가 발생할 수 있다.  
     
   sudo yum install cyrus-sasl-devel cyrus-sasl-gssapi cyrus-sasl-md5 cyrus-sasl-plain

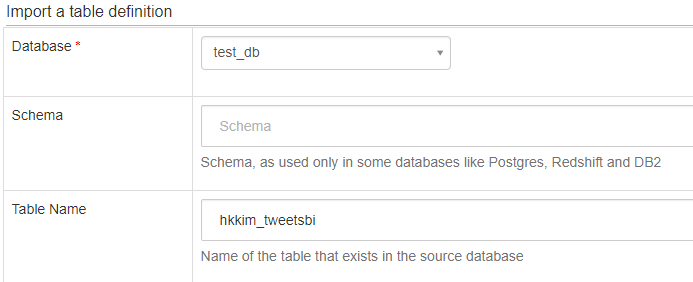
를 입력한 후, superset 프로세스를 재시작한다.   
  
ps -ef | grep superset  
kill -9 “process id”  
superset runserver  
  
**(superset stopserver, superset restartserver 등을 사용해서 서버를 재시작하고 싶었으나 적절한 명령어를 찾지 못해서, 결국 process를 죽이고 다시 시작하는 방식으로 택했다.)***출처 : https://superset.incubator.apache.org/installation.html*

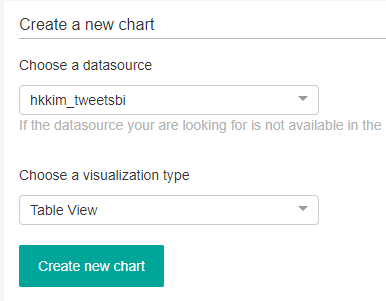
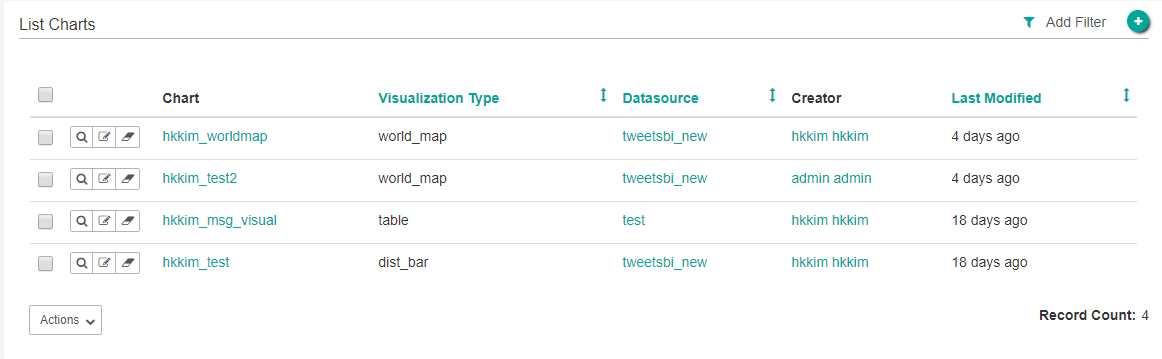
1. **Supersets 차트 생성**

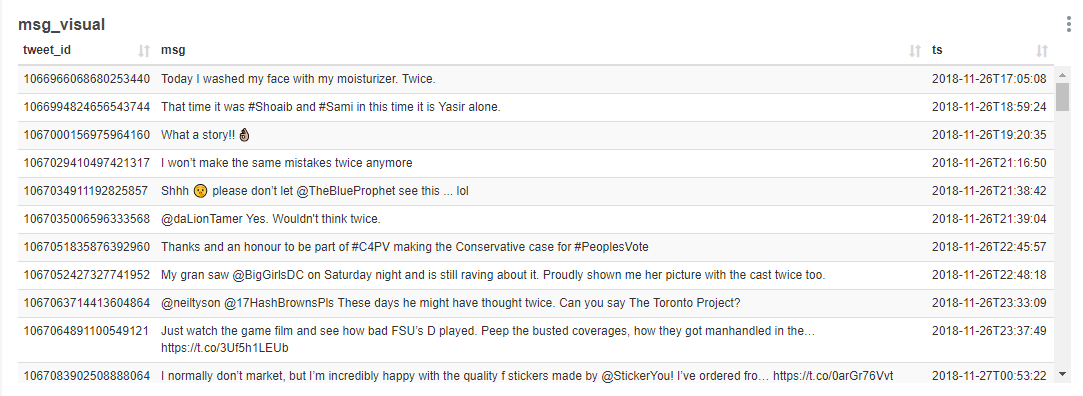
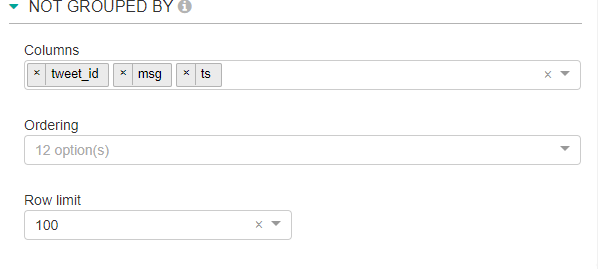
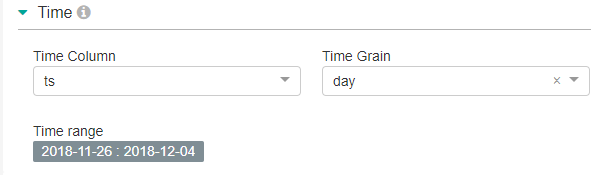
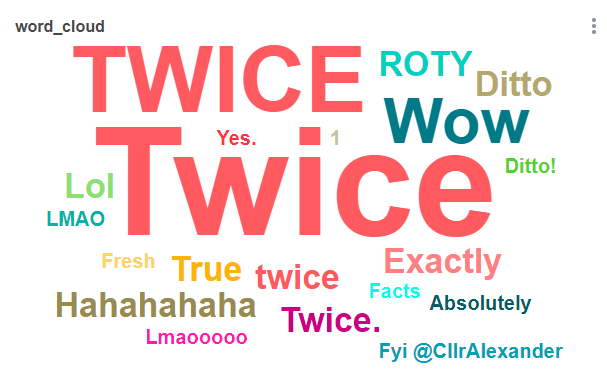
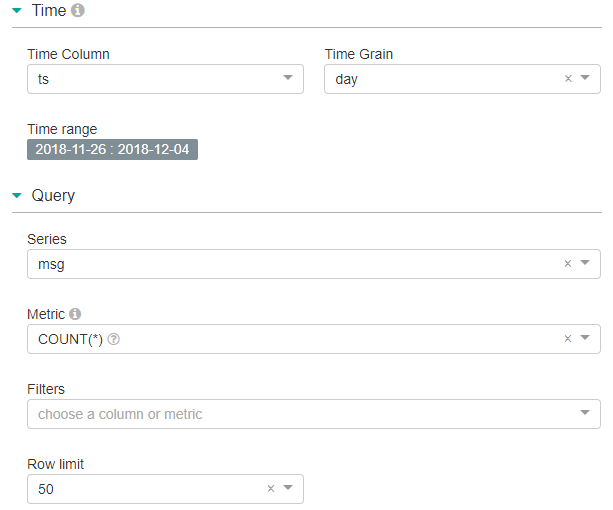
**1) 데이터 가져오기  
(1)“Source” -> “Databases”** **(2) “Add Filter” 옆의 “+” 눌러서 새 DB를 추가한다.** **(3) Database란에는 본인이 짓고 싶은 이름을, URI란에 DB주소를 적는다.** ***(Test Connection을 클릭하여 연결이 성공하면 아래와 같이 뜬다.)***  


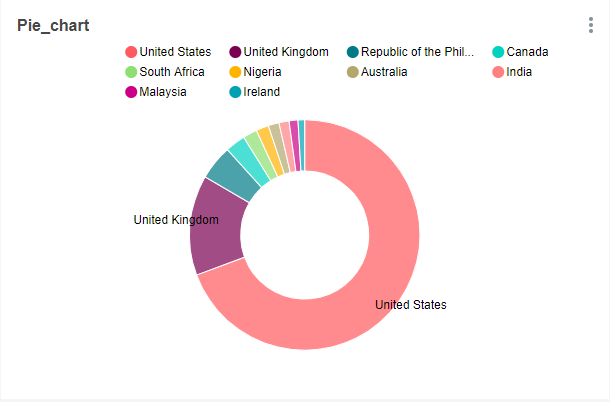
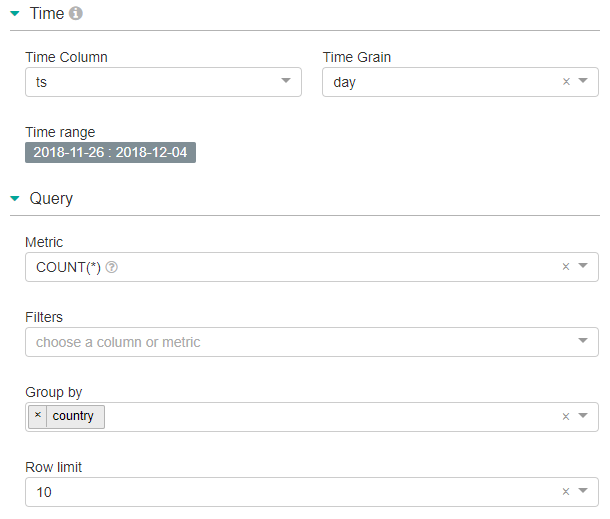
**(4) 하단에 있는****를 눌러주면 test\_db라는 이름의 DB가 생성된다.**

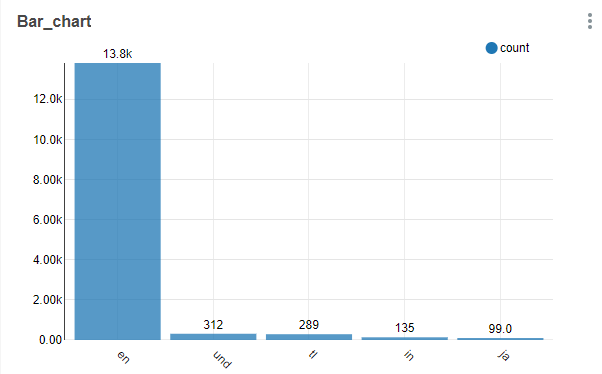
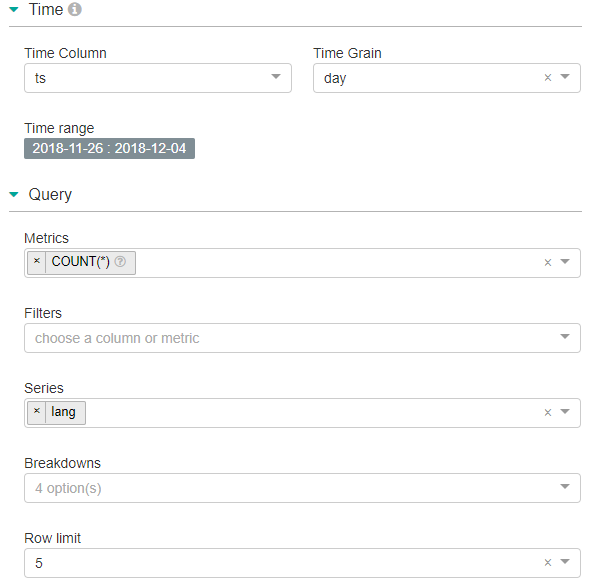
**(5) “Source” -> “Tables” -> “+”**

**(6) Database 란에는 (4)에있는 “test\_db” DB를 넣어준다.   
Table Name에는 test\_db와 연결되어 있는 hive 테이블을 하나 가져온다.**

**2) 차트 생성하기**

**2)-1 msg\_visual 차트 (“Table View” 타입 사용)** **2)-2 word\_cloud 차트 (“Word Cloud” 타입 사용)**

**2)-3 Pie\_chart 차트 (“Pie\_chart” 타입 사용)**

**2)-4 Bar\_chart 차트 (“Distrubution Bar” 타입 사용)**

**\* Dashboard의 모습**