# **How to install virtualenv:**

### **Install pip first**

sudo apt-get install python3-pip

### **Then install virtualenv using pip3**

sudo pip3 install virtualenv

### **Now create a virtual environment**

virtualenv venv

you can use any name insted of **venv**

### **You can also use a Python interpreter of your choice**

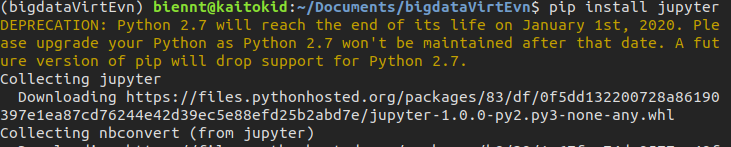
virtualenv -p /usr/bin/python2.7 venv

### **Active your virtual environment:**

source venv/bin/activate

# bigdataVirtEvn is virtualenv’s name, just like venv

### **Install jupyter in virtualenv(after active environment)**

pip install jupyter

**Download and unzip spark from link:** <http://spark.apache.org/downloads.html>

**Copy spark directory to /venv directory like: venv/spark-2.4.1-bin-hadoop2.7**

**Edit file ~./.bashrc, add the following lines:**

# Setting jupyter notebook

export PYSPARK\_DRIVER\_PYTHON=/path/to/venv/bin/jupyter

export PYSPARK\_DRIVER\_PYTHON\_OPTS='notebook'

#Setting spark

export SPARK\_HOME=/path/to/venv/spark-2.4.1-bin-hadoop2.7

export PATH=$SPARK\_HOME/bin:$PATH

export PYSPARK\_PYTHON=/path/to/venv/bin/python

#Example:

#export PYSPARK\_DRIVER\_PYTHON=/home/biennt/Documents/venv/bin/jupyter

#export PYSPARK\_DRIVER\_PYTHON\_OPTS='notebook'

#export SPARK\_HOME=/home/biennt/Documents/venv/spark-2.4.1-bin-hadoop2.7

#export PATH=$SPARK\_HOME/bin:$PATH

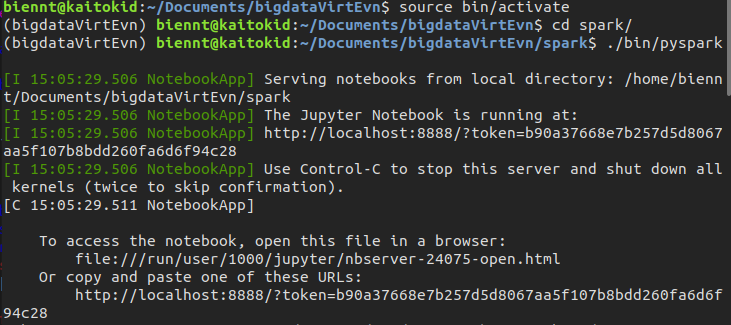
#export PYSPARK\_PYTHON=/home/biennt/Documents/venv/bin/python

Come back activated environment venv:

cd spark-2.4.1-bin-hadoop2.7

./bin/pyspark

#Click link localhost to code

****