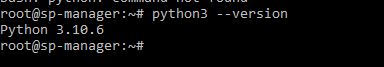
1. Get inside manager container
2. Check if python is installed. Pyspark requires at least python 3.7 or above.

$ python –version

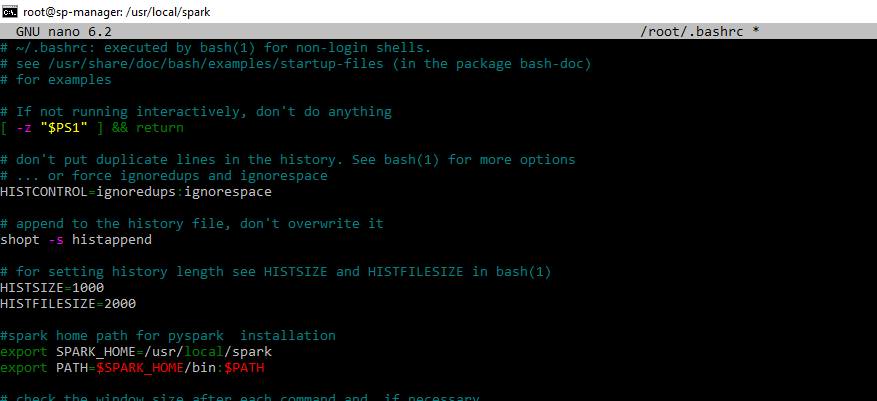


1. As we have python 3.10 , we are good to go with pyspark
2. Add spark home path in ~/.bashrc

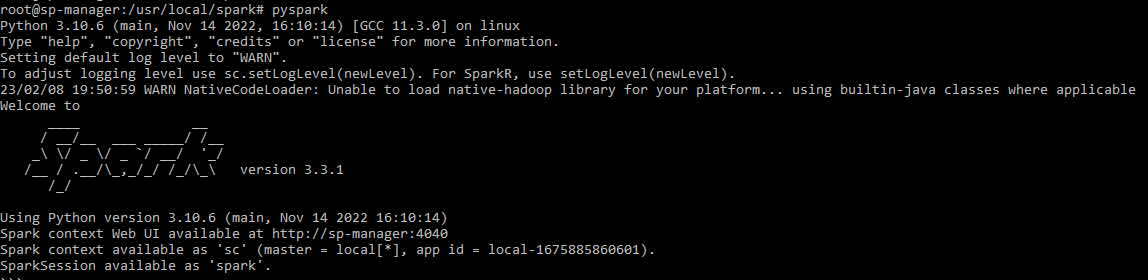
$ nano ~/.bashrc

Add below path in the file and save.

export SPARK\_HOME=/usr/local/spark  
export PATH=$SPARK\_HOME/bin:$PATH



1. Restart the terminal
2. Verify pyspark



1. Install pip and jupyter

$apt-get install pip -y && pip install jupyter -y

1. Launch Jupyter with below command so that I can be accessed from host machine on port 8888

$ jupyter notebook --ip=0.0.0.0 --port=8888 --no-browser --allow-root