

# spark 在vagrant下的搭建与启动

以前一直都使用的是vmware搭建虚拟机，但是最近听说vagrant体积小而且还方便，于是便在网上找了个虚拟系统名为 [virtualbox.box](https://www.virtualbox.org/) 用迅雷把他下载好放到一个新建的文件夹来存放这个虚拟机，然后依次输入命令来启动这个虚拟机

```
vagrant box add base virtualbox.box
vagrant init
vagrant up
```

过一会便看到这个虚拟机已经启动起来了

```
F:\box\myVagrantfile>vagrant up
Bringing machine 'default' up with 'virtualbox' provider...
==> default: Importing base box 'base'...
==> default: Matching MAC address for NAT networking...
==> default: Setting the name of the VM: myVagrantfile_default_1446443993607_98385
==> default: Clearing any previously set network interfaces...
==> default: Preparing network interfaces based on configuration...
    default: Adapter 1: nat
==> default: Forwarding ports...
    default: 22 => 2222 (adapter 1)
==> default: Booting VM...
==> default: Waiting for machine to boot. This may take a few minutes...
    default: SSH address: 127.0.0.1:2222
    default: SSH username: vagrant
    default: SSH auth method: private key
    default: Warning: Connection timeout. Retrying...
==> default: Machine booted and ready!
==> default: Checking for guest additions in VM...
    default: The guest additions on this VM do not match the installed version of
    default: VirtualBox! In most cases this is fine, but in rare cases it can
    default: prevent things such as shared folders from working properly. If you see
    default: shared folder errors, please make sure the guest additions within the
    default: virtual machine match the version of VirtualBox you have installed on
    default: your host and reload your VM.
    default:
    default: Guest Additions Version: 4.3.10
    default: VirtualBox Version: 5.0
==> default: Mounting shared folders...
    default: /vagrant => F:/box/myVagrantfile
```

这里需要注意的地方是 **SSH address, username** 一会用这个IP与用户名与密码都是 vagrant来连接该虚拟机

要记住的命令开机关机，请一定要记住要关机

vagrant up 开机

vagrant halt 关机

vagrant reload 重新启动

vagrant的配置文件vagrantfile,里可以配置虚拟机的端口等。上次因为是自己手动下载的，所以现在需要再对vagratable加上些东东

```
ipythonPort = 8001 #放到config外面
config.vm.network :forwarded_port, host: ipythonPort, guest:
ipythonPort, auto_correct: true
config.vm.network :forwarded_port, host: 4040, guest: 4040,
auto_correct: true
config.vm.hostname = "sparkvm"
config.vm.usable_port_range = 4040..4090
```

现在就开始学习吧

What would Fischer say is a potential flaw of the Seven Countries Study?

Correlation does not imply causation

What other lessons can a data scientist take away from the Seven Countries Study?

Use more data

Explore more types of data/factors