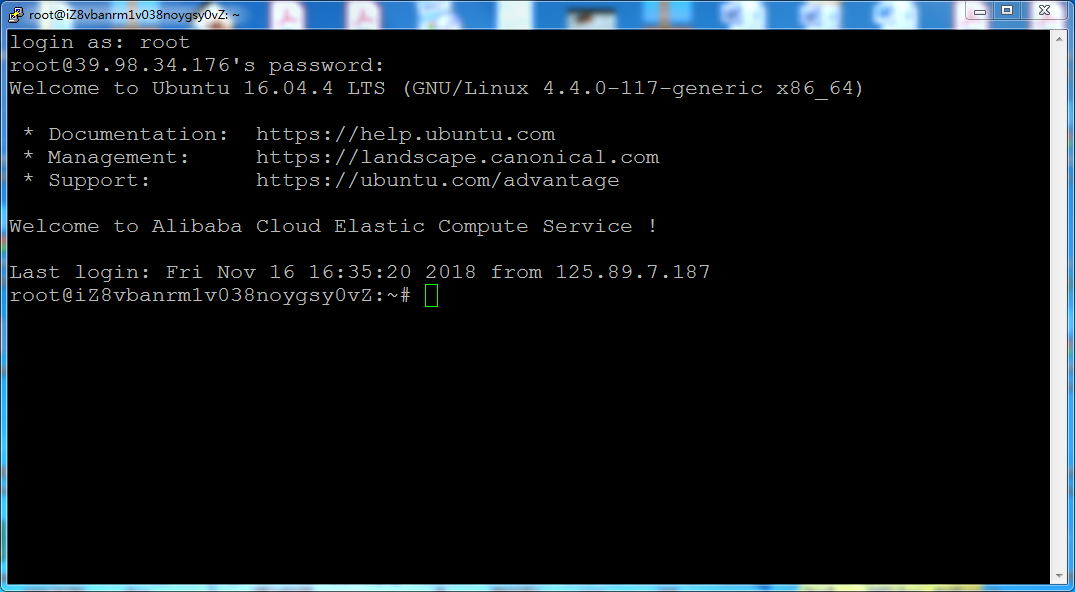
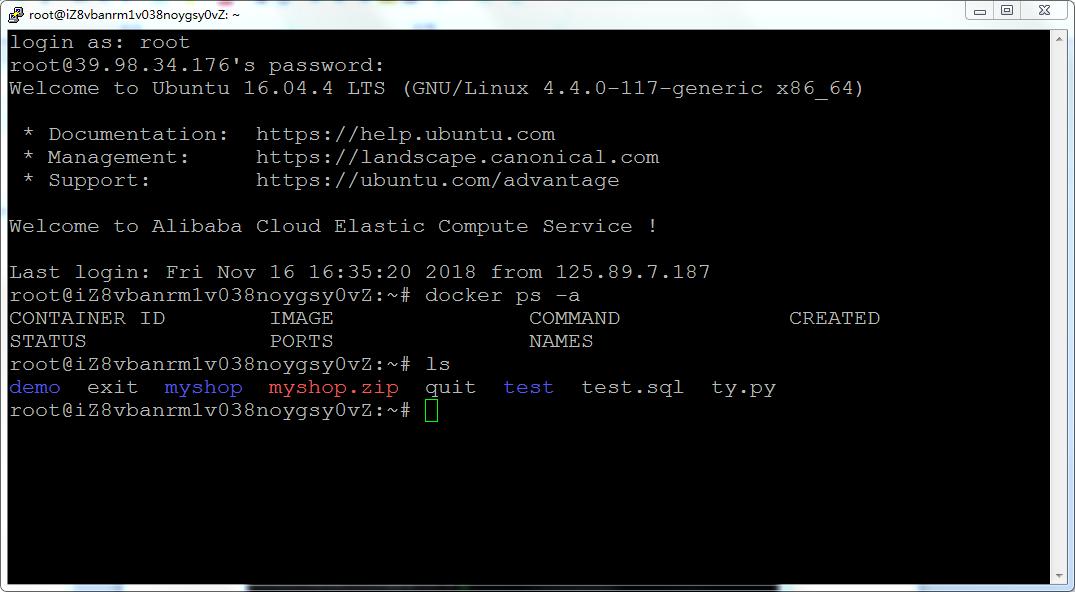
Docker制作爬虫镜像案例

1. 登录远程服务器

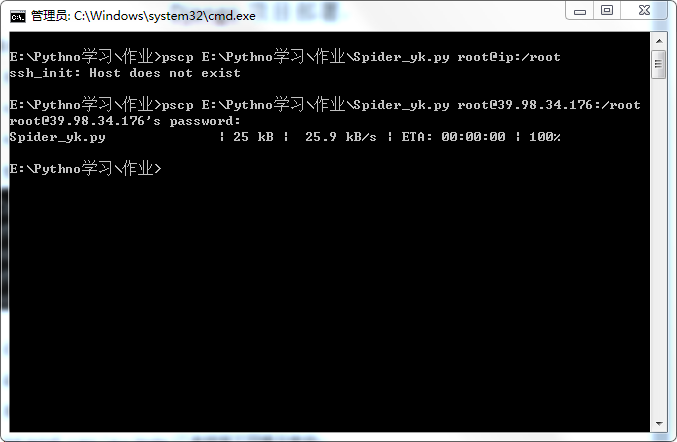


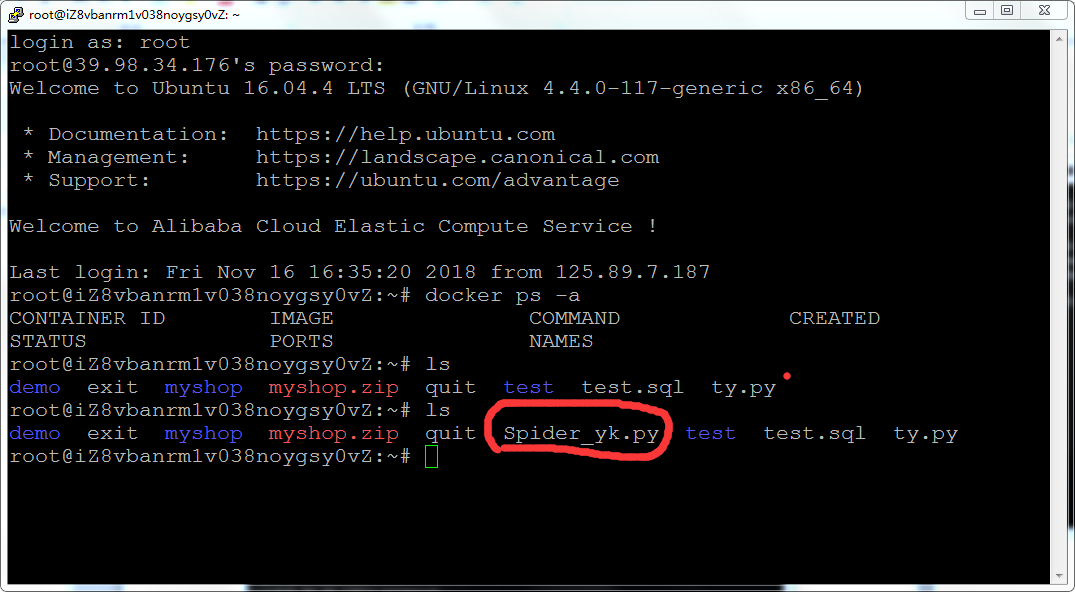
1. 查看当前是否有运行的镜像，同时是否存在Dockerfil镜像文件。



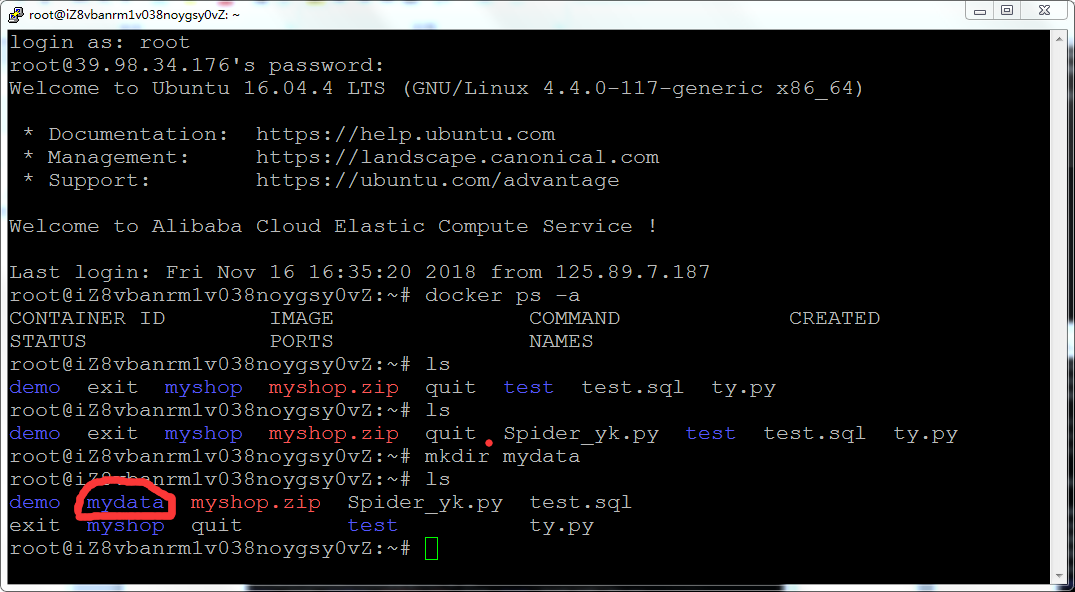
1. 因为当前没有镜像运行，也没有自定义的镜像文件，那么我先创建一个文件夹mydata并在里面上传一个叫Spider\_yk.py的文件，该文件为一个爬虫文件主要是爬取优酷电影信息。而且创建一个自己的镜像文件Dockerfil,如图详解。

（1）、上传文件

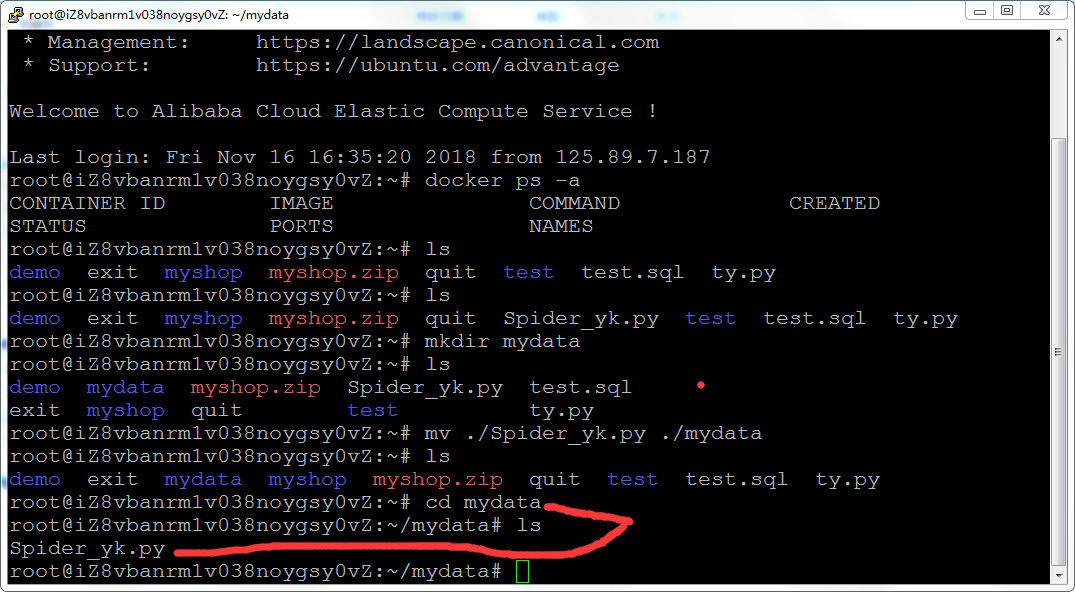




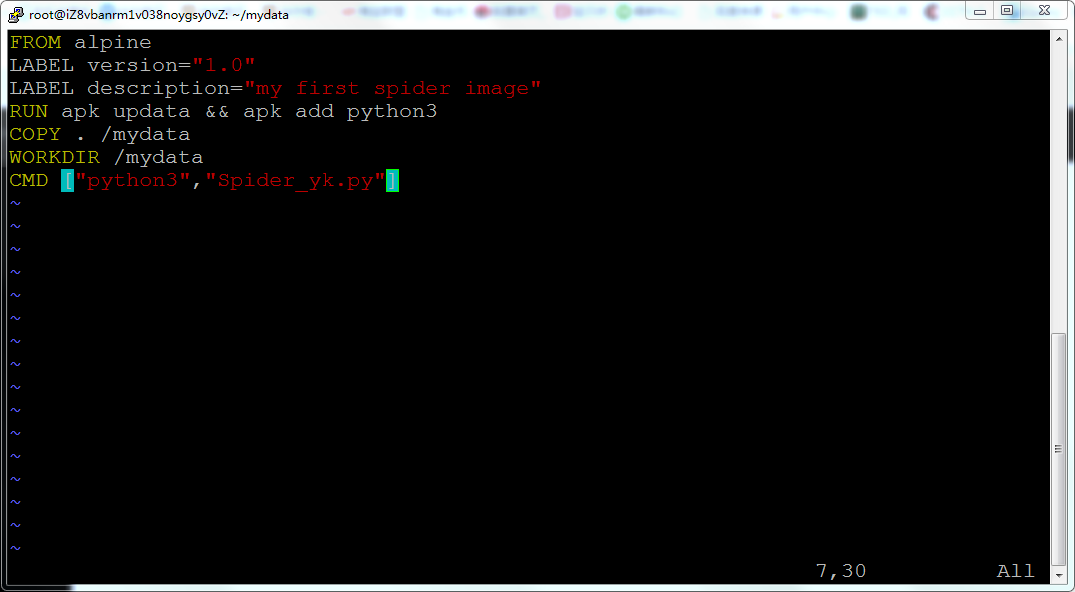
创建文件夹mydata



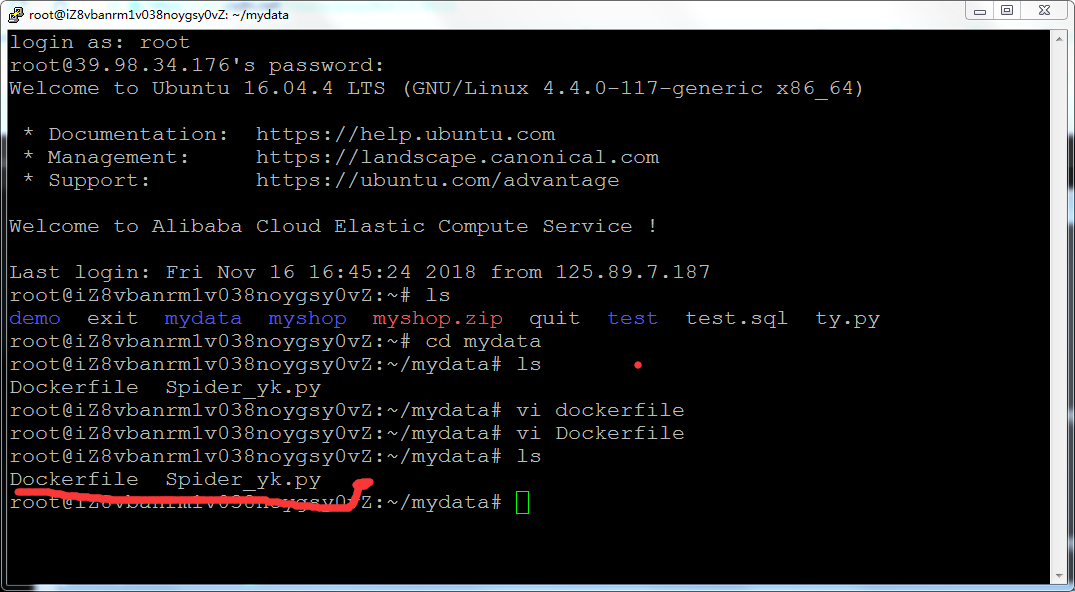
把文件移动进入mydata



开始制作镜像文件Dockerfile

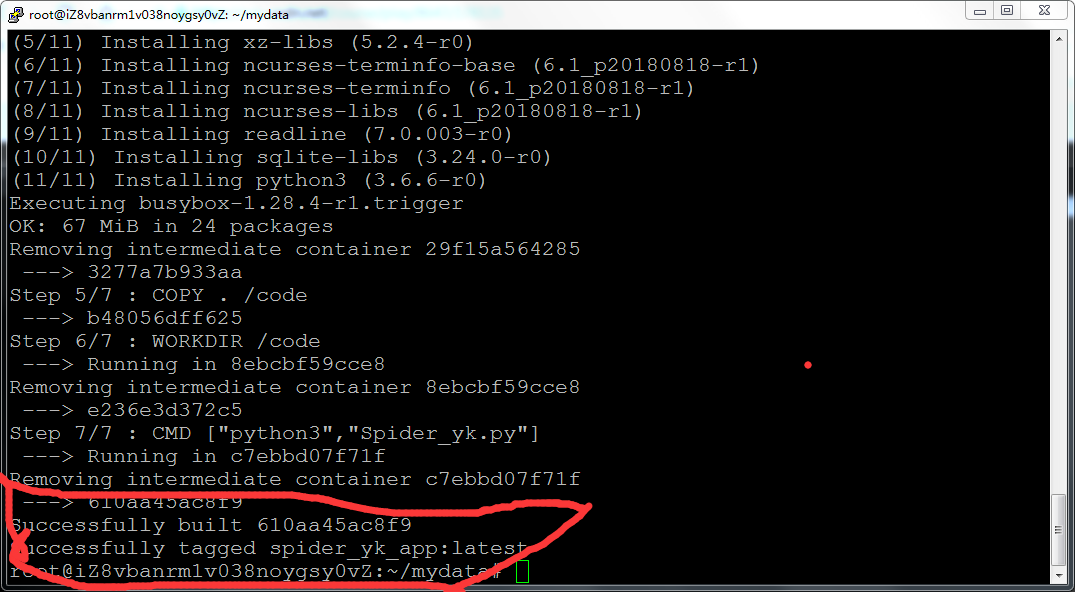


执行发现错误修改上图中的updata变成update

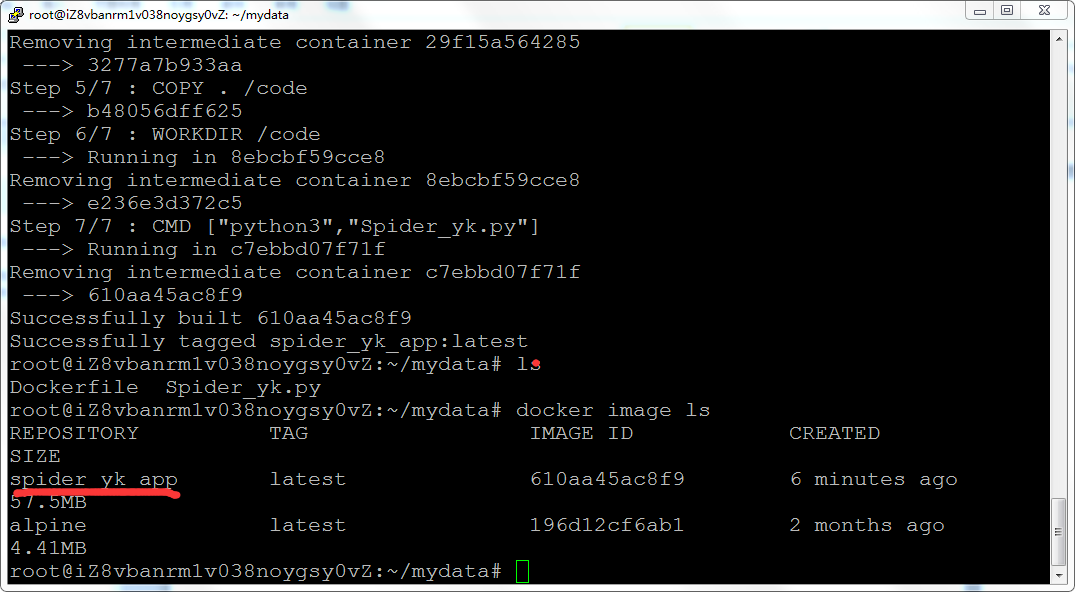


文件齐全了，可以创建镜像

Docker build –t spider\_yk\_app .

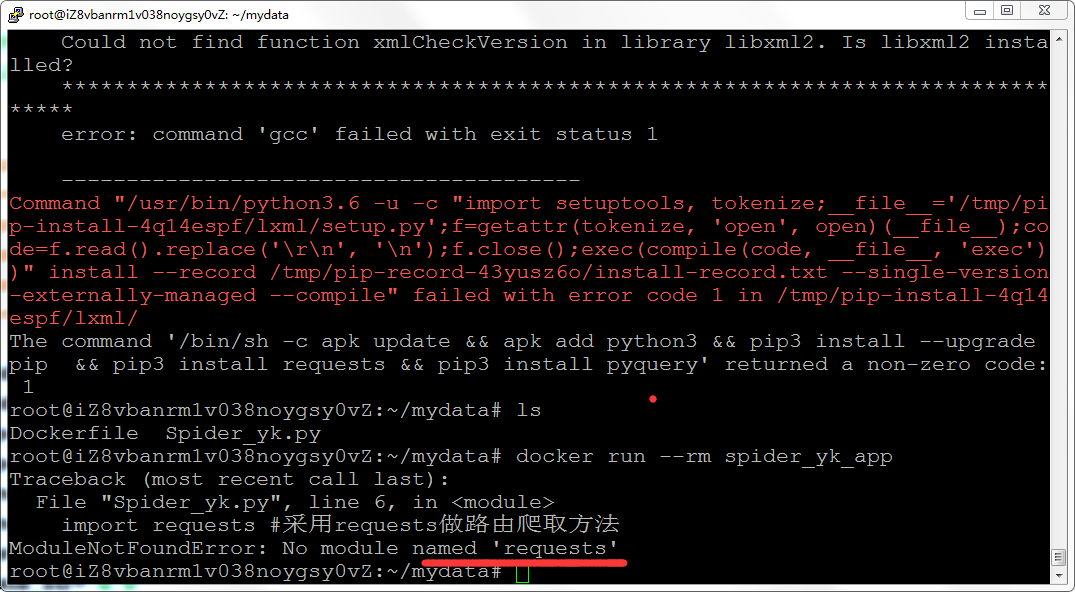


代表成功制作到了镜像文件。

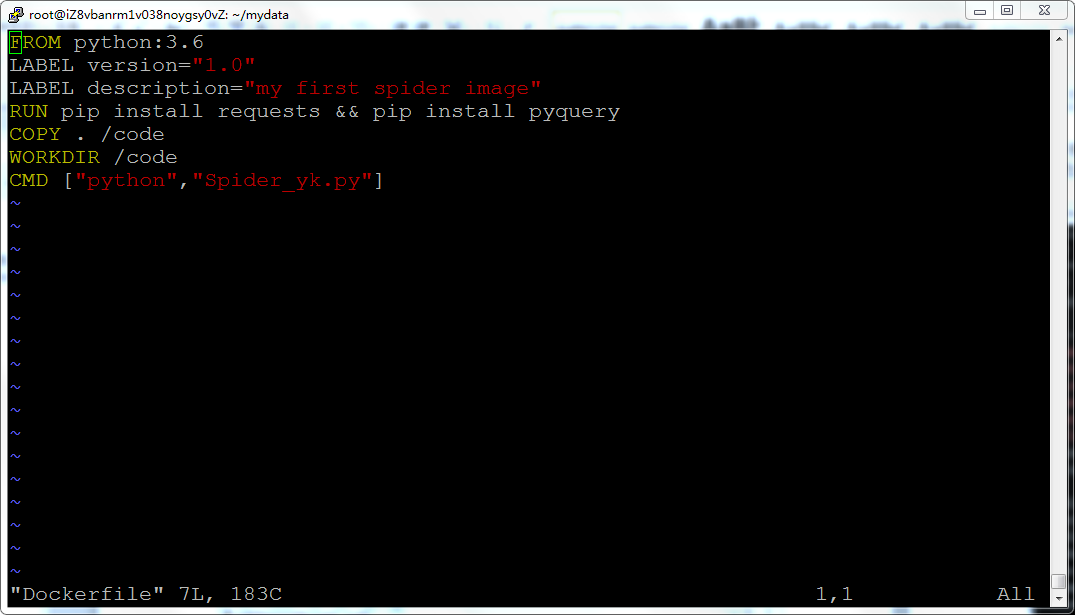


最后只要执行镜像文件即可

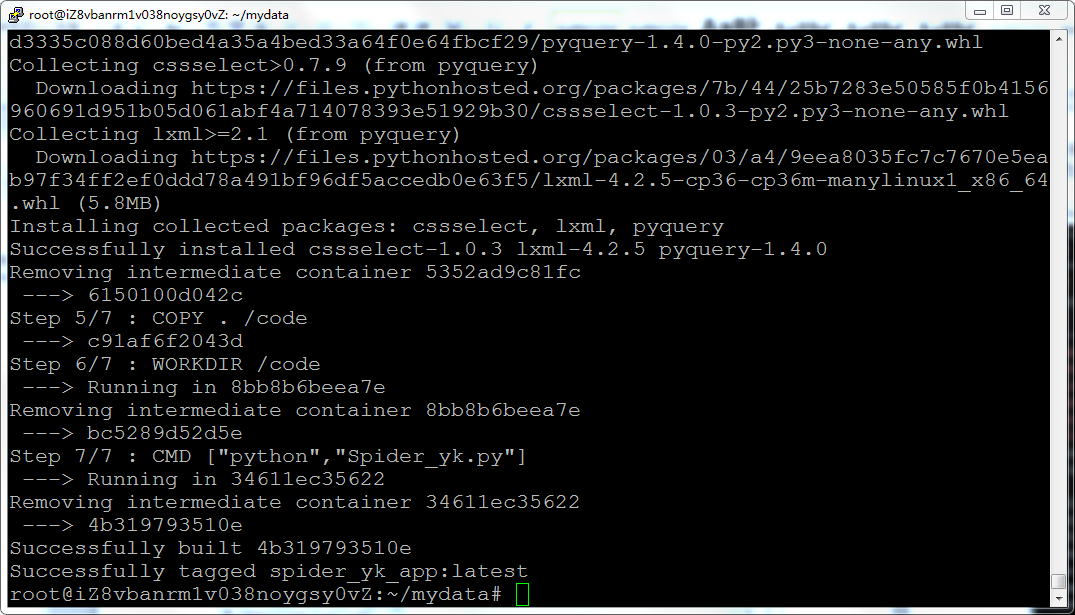
执行是出现错误，发现写的爬虫文件里面有要依赖的库函数，因此必须重新更改Dockerfile文件



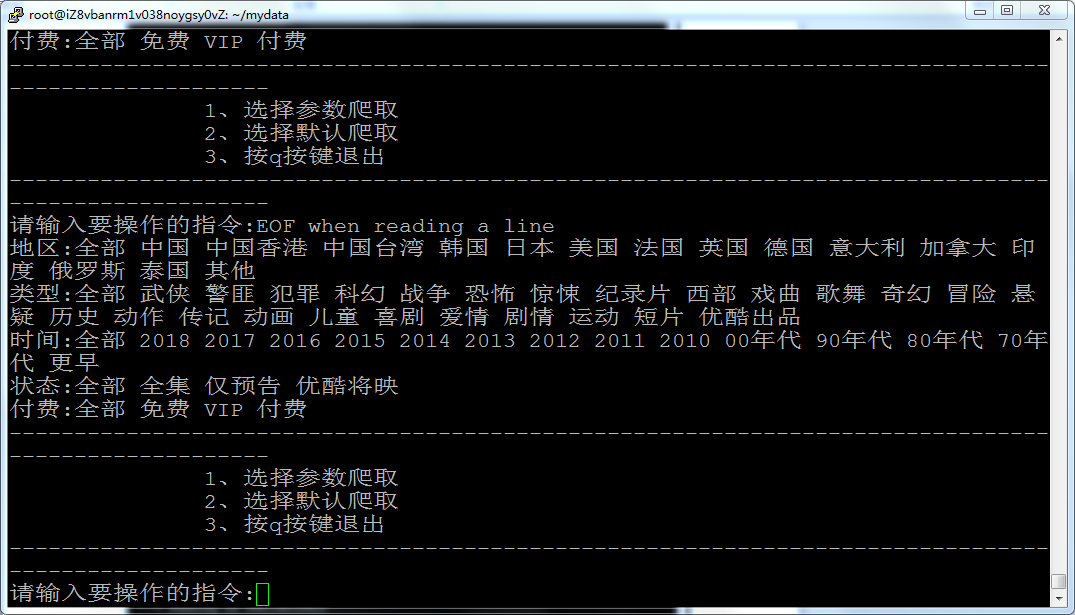
后进行更改还是依旧，通过咨询同学后得出目前同学都没有通过alpine和ubuntu下进行镜像，必须改成python3.6；这个时候我突然有点感觉，后面再测试；先按同学们建议改成如下

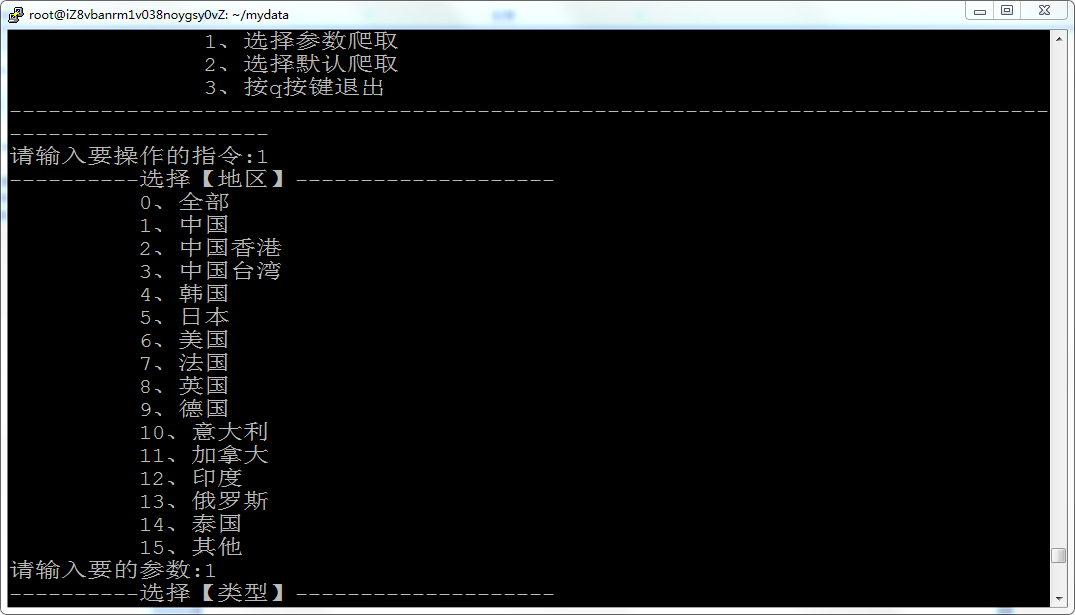


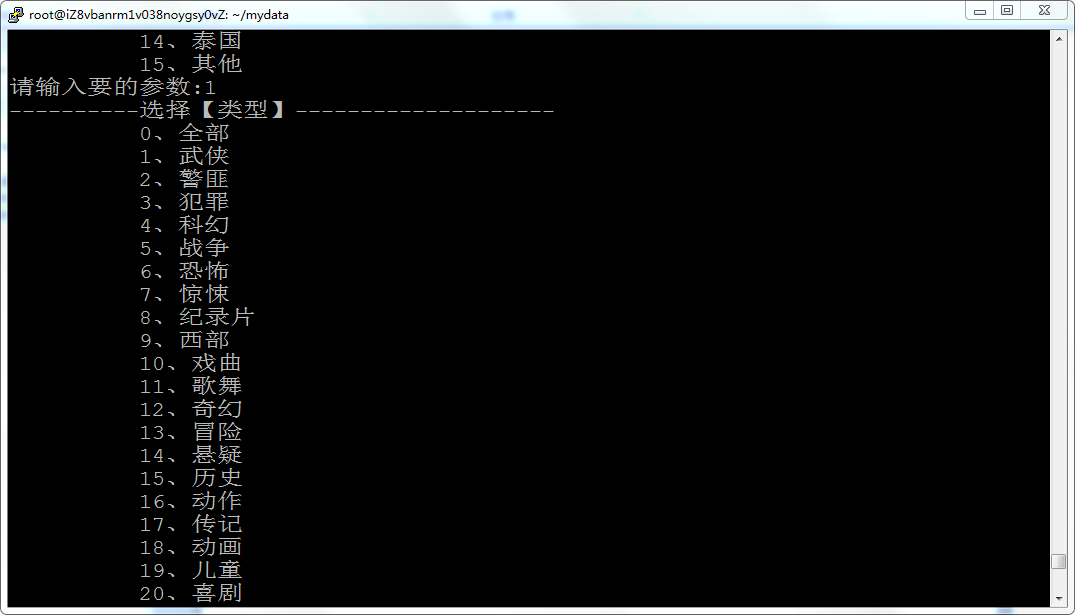
镜像制作果真成功（果然是过来人）

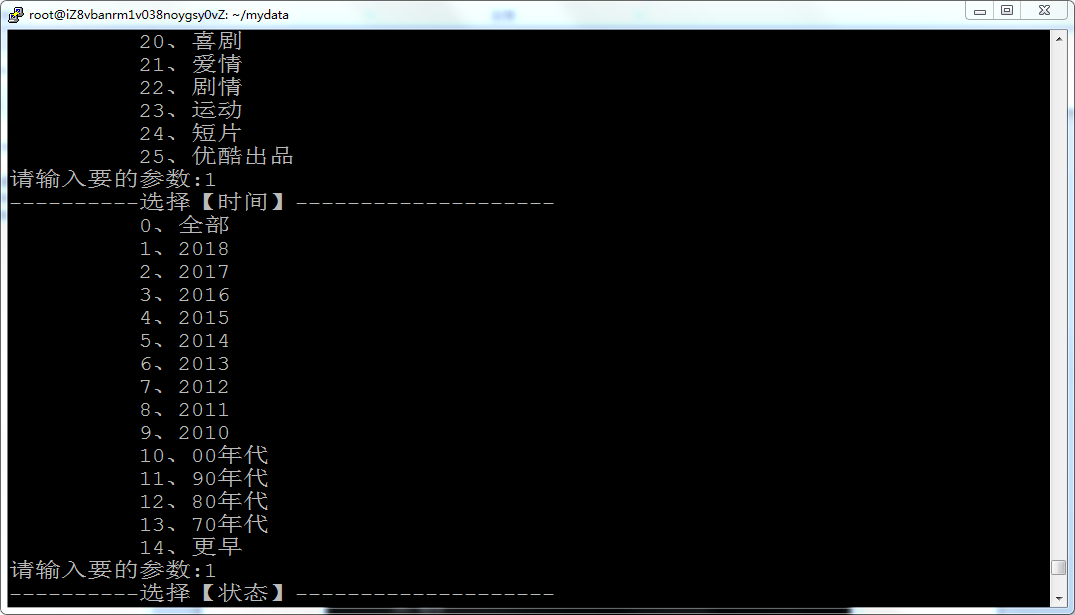


然后运行镜像(成功出现)



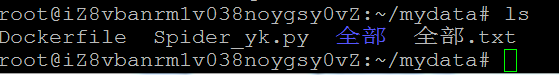








程序就可以爬取优酷网的任何电影信息，同时保存在当前目录



因为选择默认为全部，否则选择什么类型就是类型的名称，一个是图片保存文件将，一个是电影的信息.txt文档。