《基于mapreduce技术的新冠肺炎数据分析报告》

1. 环境准备
2. dokcer容器

docker run -it --privileged --name="hadoop1" --hostname="hadoop1" -v /Users/yuanxingyou/hadoop/:/hadoop -p 9870:9870 -p 50070:50070 -p 9000:9000 -p 3306:3306 d8609b6a89df

2.在etc/profile下添加

export PATH=$PATH:/hadoop/hadoop-3.3.0/sbin:/hadoop/jdk1.8.0\_281/bin:hadoop/jdk1.8.0\_281/bin:$PATH

3.在hadoop-3.3.0/etc/hadoop-env.sh修改

export JAVA\_HOME=/hadoop/jdk1.8.0\_281

1. ssh免密认证

ssh-keygen -t rsa

1. 配置hive
2. 实现：
3. 创建data\_process.py文件对数据CSV文件提取湿度，日期，省份
4. 用hive语言进行查询，计数

三．主要代码：

1.data\_process.py代码



1. hive语句
2. 结果:
3. 运行截图：

