

win10系统手把手教你docker练习场满分

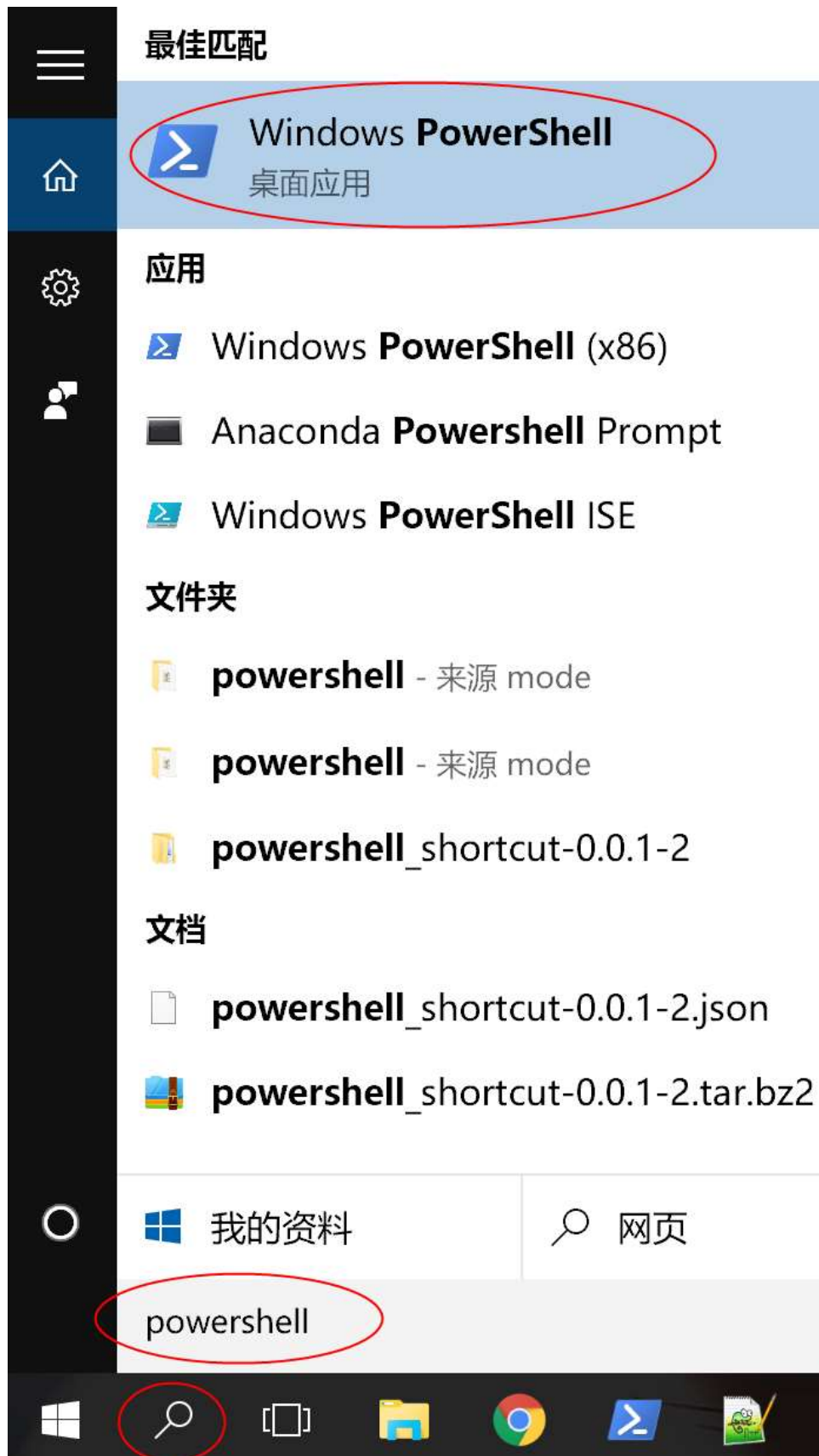
一、配置docker环境

1. 推荐下载Docker Desktop

下载地址：<https://docs.docker.com/docker-for-windows/install/>

你以为点开网址就开始下载了？别急！

先以各种方式打开powershell，这里推荐一种最无脑的方法：电脑左下角点搜索powershell



打开以后，输入如下命令（可直接一起复制粘贴），查看自己的电脑版本和 buildnumber

```
Get-WmiObject -Class Win32_OperatingSystem | % Caption
```

```
Get-WmiObject -Class Win32_OperatingSystem | % Buildnumber
```

接下来查看自己电脑，是否符合页面上最新版本的Docker Desktop的要求



如果电脑不符合要求：

方法一：升级电脑系统

方法二：打开此网页，下载符合要求的历史版本

<https://docs.docker.com/docker-for-windows/release-notes/>

2. 配置 Docker Desktop

在系统右下角托盘图标内右键菜单选择 Settings，打开配置窗口后查看左侧导航菜单：

- 1) 选择 Shared Drives（不同版本菜单名字可能会稍微有点变化，中心意思不变），根据需求选择C盘、D盘、E盘等等。
- 2) 选择 Advanced（不同版本菜单名字可能会稍微有点变化，中心意思不变），根据需求修改images等路径（最好不要存在C盘，文件会越来越大）
- 3) 使用镜像加速器：选择 Daemon。旧版本在Registry mirrors里直接添加 <https://5w5jh673.mirror.aliyuncs.com> 网址；新版本选择 Docker Daemon。编辑窗口内的JSON串，填写下方加速器地址：

```
{
  "registry-mirrors": ["https://5w5jh673.mirror.aliyuncs.com"]
}
```

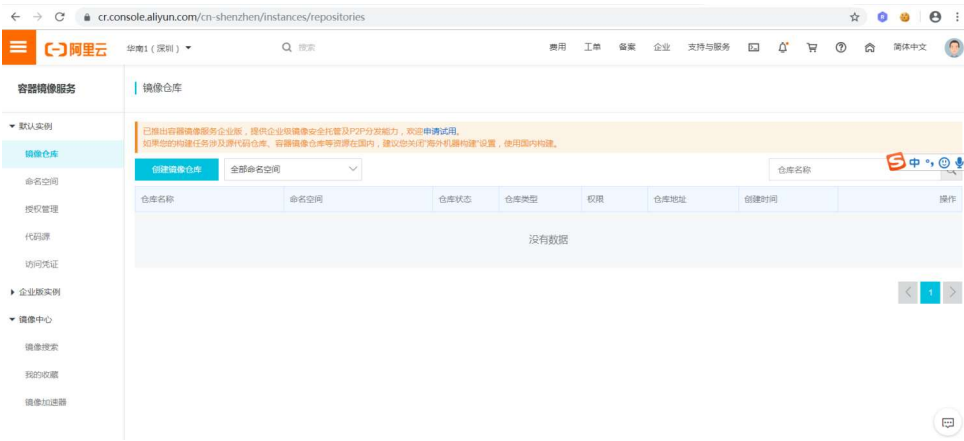
二、开通容器镜像服务（推荐阿里云）

☆☆为方便懒癌同学，这里做搬运工，把官方教学搬过来，如下：

（官方教学见网址

<https://tianchi.aliyun.com/competition/entrance/231759/tab/226>)

阿里云容器服务地址为 (<https://cr.console.aliyun.com>)



切换标签页到命名空间，创建地址唯一的命名空间



根据任务/比赛要求选择对应的地域，其他的按照自己需求选择或填写。

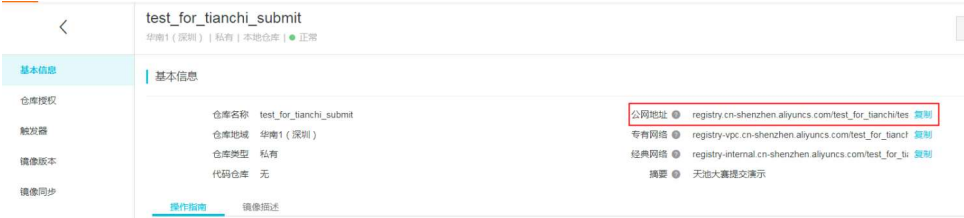


下一步，选择本地仓库，不建议其他选项，完成创建。



点击管理，可查看详情。





三、编写文件

1.读题:

参与者可分阶段提交容器镜像完成以下3个任务（分数依次占 30/30/40） , 根据评分系统的分数返回验证任务的完成情况。

- 输出Hello world
- 计算 /tcdata/num_list.csv中一列数字的总和
- 在/tcdata/num_list.csv文件中寻找最大的10个数，从大到小生成一个List

num_list.csv文件中只有一列不为负的整数，其中存在重复值，示例如下：
生成入口脚本 [run.sh](#)，放置于镜像工作目录。运行后生成结果result.json放置于工作目录（与run.sh同目录），评分系统将根据result.json进行打分。json文件如下所示：

```
{
  "Q1": "Hello world",
  "Q2": sum值,
  "Q3": [top10_list]
}
```

2.操作

2.1 新建文件夹，如：`docker_submit_demo`（名字可自拟），并在此文件夹下，新建`tcdata`文件夹

2.2 利用python在`docker_submit_demo`文件夹下编写文件（也可用其他工具），图片和代码如下：

名称	修改日期	类型
tcdata	2021/1/22 17:50	文件夹
Dockerfile	2021/1/22 18:16	文件
hello_world.ipynb	2021/1/22 10:52	IPYNB 文件
hello_world.py	2021/1/22 18:13	PY 文件
requirements.txt	2021/1/22 13:37	Notepad++ Docu...
result.json	2021/1/22 17:50	JSON 文件
run.sh	2021/1/20 19:31	Shell Script

☆☆为了方便演示，首先用jupyter notebook写一遍，最后放py文件整合代码，懒癌同学可直接跳过

```
import pandas as pd
import numpy as np
import json
```

#num_list.csv文件在天池给的基础镜像中，已在tcdata文件夹中包含
#这里为了方便测试程序，自己编写一个num_list.csv文件，实际提交结果时，用不到此文件

```
data=np.random.randint(1,100,200)
```

```
data
array([[75, 29, 95, 79, 70, 40, 14, 6, 89, 95, 8, 4, 58, 9, 39, 16, 26,
        25, 31, 78, 98, 33, 5, 98, 20, 64, 27, 72, 85, 71, 74, 54, 55, 39,
        5, 98, 66, 48, 78, 33, 94, 97, 23, 37, 9, 48, 40, 88, 76, 25, 88,
        62, 29, 34, 80, 81, 48, 61, 64, 39, 78, 84, 36, 96, 10, 36, 74, 85,
        34, 82, 14, 30, 14, 79, 97, 77, 71, 28, 26, 1, 98, 67, 58, 34, 46,
        57, 99, 43, 98, 88, 5, 5, 72, 63, 29, 53, 62, 77, 47, 80, 96, 79,
        22, 27, 93, 38, 45, 75, 19, 94, 92, 61, 29, 17, 63, 36, 30, 94, 83,
        47, 4, 23, 38, 80, 39, 99, 67, 2, 41, 20, 53, 45, 90, 1, 85, 18,
        79, 64, 75, 89, 1, 65, 93, 27, 7, 25, 31, 92, 30, 50, 49, 10, 2,
        45, 87, 73, 68, 35, 14, 15, 30, 21, 80, 9, 89, 50, 16, 61, 22, 85,
        87, 44, 42, 63, 74, 53, 24, 92, 49, 25, 67, 22, 17, 47, 58, 64, 78,
        44, 92, 44, 26, 5, 4, 75, 95, 10, 79, 34, 6, 49])
```

```
data=pd.DataFrame(data)
```

data

	0
0	75
1	29
2	95
3	79
4	70
...	...
195	10
196	79
197	34
198	6
199	49

200 rows × 1 columns

```
data.to_csv("./tcdata/num_list.csv",index=False,header=False)
```

```
data=pd.read_csv("./tcdata/num_list.csv",header=None)
```

data

	0
0	75
1	29
2	95
3	79
4	70
...	...
195	10
196	79
197	34
198	6
199	49

200 rows × 1 columns

```

#第一题
result_1="Hello world"

#第二题
result_2=0
for i,num in enumerate(data[0]):
    result_2+=num

#第三题
data.sort_values(by=0,ascending=False,inplace=True)
result_3=data[0][:10]

result_3
125    99
86     99
35     98
80     98
88     98
23     98
20     98
41     97
74     97
100    96
Name: 0, dtype: int64

result_3=list(result_3)
result_3
[99, 99, 98, 98, 98, 98, 97, 97, 96]

result={"Q1":result_1,
        "Q2":result_2,
        "Q3":result_3
        }

with open('result.json', 'w', encoding='utf-8') as f:
    json.dump(result, f)

```

☆☆代码整合如下

```

#!/usr/bin/env python
# coding: utf-8
import pandas as pd
import numpy as np
import json

#data=np.random.randint(1,100,200)
#data=pd.DataFrame(data)
#data.to_csv("./tcdata/num_list.csv",index=False,header=False)
data=pd.read_csv("/tcdata/num_list.csv",header=None)#天池python镜像默认包含此文件，自己测试用如
#data=pd.read_csv("./tcdata/num_list.csv",header=None)

#第一题
result_1="Hello world"

#第二题
result_2=0
for i,num in enumerate(data[0]):
    result_2+=num

```

```
#第三题
data.sort_values(by=0,ascending=False,inplace=True)
result_3=data[0][:10]
result_3=list(result_3)

result={"Q1":result_1,
        "Q2":result_2,
        "Q3":result_3
        }

with open('result.json', 'w', encoding='utf-8') as f:
    json.dump(result, f)
```

2.3 由于使用python的时候用到了一些库，所以需要编写requirements.txt文件：

方法一：在python里使用pip freeze > requirements.txt（此方法缺点：会把许多我们此次文件用不到的库也写进requirements.txt里。当然，如果用此方法，你可以自己打开生成文件，删减掉不用的库）

方法二：使用 pipreqs，github地址为：<https://github.com/bndr/pipreqs> 代码如下：

```
# 安装
pip install pipreqs
# 在当前目录生成
pipreqs . --encoding=utf8 --force
requirements.txt文件代码如下：
```

```
pandas==0.25.1
numpy==1.19.4
```

2.4 编写run.sh文件（有多种方法创建run.sh文件，最简单的方法就是用文本编辑器生成一个.txt文件，把后缀改成.sh）

用文本编辑器编辑（推荐Notepad++），代码如下：

```
#!/bin/sh
CURDIR="`dirname $0`" #获取此脚本所在目录
echo $CURDIR
cd $CURDIR #切换到该脚本所在目录
python hello_world.py
```

- 如果直接在hello_world.py文件所在位置运行指令sh run.sh，上述代码可以简化为python hello_world.py
- 如果系统装了python2和python3，那么要改为python3 hello_world.py
- 运行sh run.sh推荐使用Git Bash

2.5 编写Dockerfile文件（有多种方法创建Dockerfile文件，最简单的方法就是用文本编辑器生成一个.txt文件，把后缀去掉）：


```
# Base Images
## 从天池基础镜像构建
FROM registry.cn-shanghai.aliyuncs.com/tcc-public/python:3

## 把当前文件夹里的文件构建到镜像的根目录下（.后面有空格，不能直接跟/）
ADD . /

## 指定默认工作目录为根目录（需要把run.sh和生成的结果文件都放在该文件夹下，提交后才能运行）
WORKDIR /

## Install Requirements (requirements.txt包含python包的版本)
## 这里使用清华镜像加速安装
RUN pip install -i https://pypi.tuna.tsinghua.edu.cn/simple -r requirements.txt

## 镜像启动后统一执行 sh run.sh
CMD ["sh", "run.sh"]
☆☆Dockerfile指令介绍传送门 https://yq.aliyun.com/articles/735190?spm=5176.12281978.0.0.37722232Z7TbgD
```

在docker_submit_demo文件夹下，地址栏内输入powershell，打开powershell，进行指令输入

名称	修改日期	类型
tcddata	2021/1/22 17:50	文件夹
Dockerfile	2021/1/22 18:16	文件
hello_world.ipynb	2021/1/22 10:52	IPYNB 文件
hello_world.py	2021/1/22 18:13	PY 文件
requirements.txt	2021/1/22 13:37	Notepad++ Docu...
result.json	2021/1/22 17:50	JSON 文件
run.sh	2021/1/20 19:31	Shell Script

```
## docker登陆（push镜像需要登陆,服务器地址为赛事要求的，这里以上海为例）：
docker login --username=你自己的用户名 registry.cn-shanghai.aliyuncs.com
```

```
## 不要忘记加.
## 公网地址上述已经提到过
docker build -t 公网地址/版本号 .
```

```
## 构建完成后可先验证是否正常运行，正常运行后再进行推送。
CPU镜像: docker run 公网地址/版本号 sh run.sh
GPU镜像: nvidia-docker run 公网地址/版本号 sh run.sh
```

```
## 推送到镜像仓库,如果出错，可能没登录
docker push 公网地址/版本号
附：
```

四、提交结果

(官方介绍传送门

<https://tianchi.aliyun.com/competition/entrance/231759/tab/226>)

☆☆提交结果界面的镜像路径：公网地址/版本号

结果如下：

第一期	
	<div><div></div><div>日期: 2021-01-22 18:38:07 排名: 无 查看日志</div><div>score: 100</div></div>

*****完结散花*****

附件submit.zip为提交结果的代码，可参考