**数据科学导论**

**Lab1：Python入门**

**杭州电子科技大学**

**2023年秋**

目录

[1. Windows环境Anaconda3安装和适用 1](#_Toc1516357187)

[1.1. 下载Anaconda3安装包 1](#_Toc1226802960)

[1.2. 安装Anaconda3安装包（windows为例） 2](#_Toc886853873)

[1.3. 验证是否成功安装 7](#_Toc1816533331)

[1.4. 使用Jupyter notebook 9](#_Toc1848168365)

[2. 实验一、使用基本数据类型 10](#_Toc962240347)

[2.1. 实验目的 10](#_Toc1821650119)

[2.2. 实验内容 10](#_Toc1946678401)

[2.3. 实验报告要求 12](#_Toc910523562)

[3. 实验二、程序控制结构 12](#_Toc201038012)

[3.1. 实验目的 12](#_Toc854090953)

[3.2. 实验内容 12](#_Toc925950523)

[3.3. 实验报告要求 13](#_Toc1783933899)

[4. 实验三、股价可视化 13](#_Toc1557844726)

[4.1. 实验目的 13](#_Toc575685658)

[4.2. 实验内容 13](#_Toc1135024271)

[4.3. 实验报告要求 15](#_Toc255686396)

[5. 实验报告 15](#_Toc206479925)

**Python入门实验**

本实验用以学习Python基础知识，要求学生通过3个实验基本掌握Python的编程方法。上机报告最迟在11月5日（星期日）晚12点前发送至https://www.wjx.top/vm/t5wvMXw.aspx# 。上机报告的内容应包括相关代码（xx.ipynb）和这份报告（屏幕截图），标题为：[学号][姓名]-Python基础。

实验时数：3学时。



# Windows环境Anaconda3安装和适用

Anaconda是一个软件包管理器、一个环境管理器以及一个Python发行版，其中包含许多开源软件包的集合(Numpy，Scikit-learn，Scipy，Pandas等)。如果在安装Anaconda之后需要其他软件包，则可以使用Anaconda的软件包管理器、conda或pip安装这些软件包。这是非常有利的，因为您不必自己管理多个软件包之间的依赖关系。Conda甚至使在Python 2和3之间切换变得容易。实际上，安装Anaconda也是安装Jupyter Notebooks的常见方法。

## 下载Anaconda3安装包

Anaconda官网网站：[Anaconda | Individual Edition](https://www.anaconda.com/products/individual)下载安装包速度比较慢，国内同学不建议尝试。我们可以使用清华大学开源软件镜像站：[Index of /anaconda/archive/ | 清华大学开源软件镜像站 | Tsinghua Open Source Mirror](https://mirrors.tuna.tsinghua.edu.cn/anaconda/archive/)，请选择适合自己系统自行下载（图1）。为方便大家实验，已提供下载好的Anaconda软件。

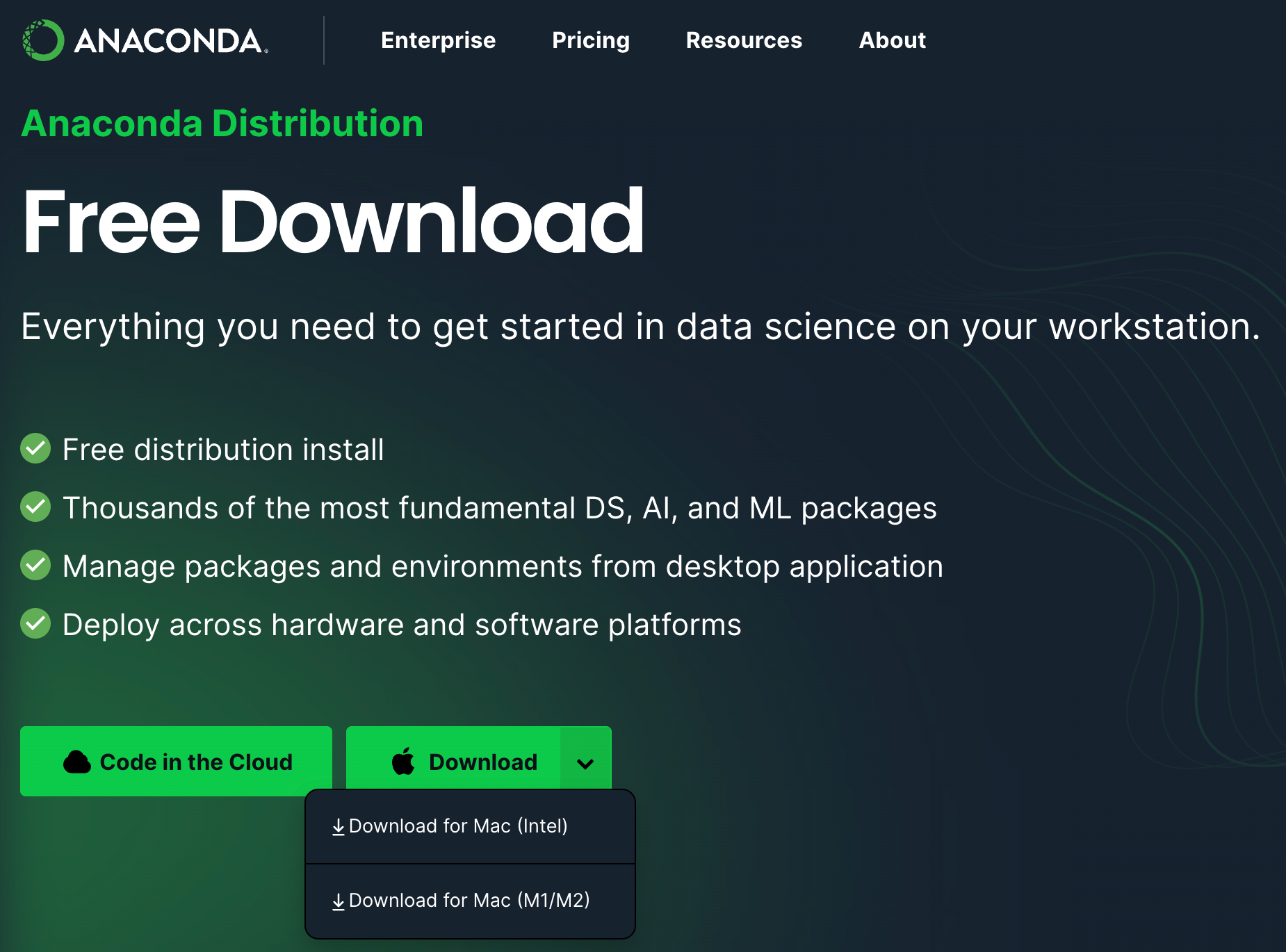


图1：下载Anaconda安装包

下载完成后（windows）如图2：

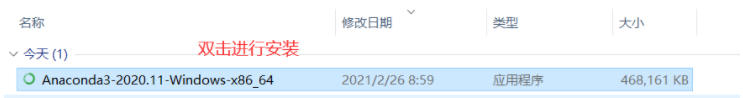


图2：成功下载Anaconda安装包

## 安装Anaconda3安装包（windows为例）

打开安装包，点击Next。

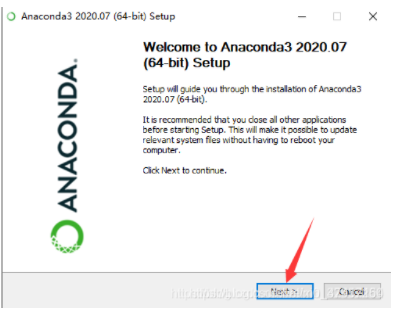


图3：打开Anaconda安装

点击I Agree。

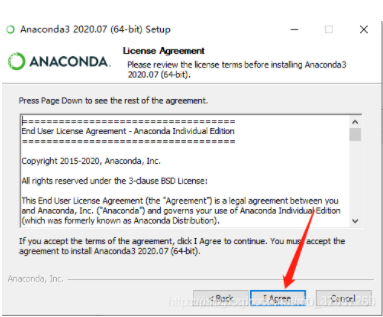


图4：安装Anaconda

Just me还是All Users？假如你的电脑有好几个 Users ，才需要考虑这个问题.其实我们电脑一般就一个 User，就我们一个人使用，如果你的电脑有多个用户，选择All Users。这里直接 All User继续点击 Next。

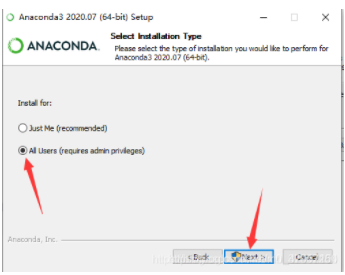


图5：安装Anaconda

Destination Folder 是“目标文件夹”的意思，可以选择安装到什么地方。默认是安装C:\ProgramData\Anaconda3文件夹下。你也可以选择 Browse… ，选择想要安装的文件夹。继续点击 Next。

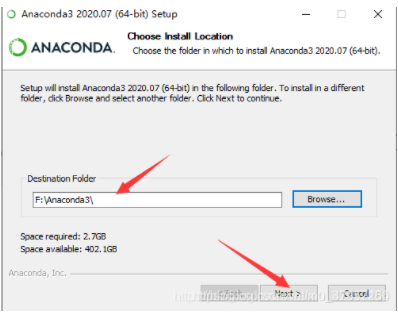


图6：安装Anaconda

第一个选项是加入环境变量，建议同学们选中这个选项，否则后面需要自己配置环境变量比较麻烦，第二个是默认使用 Python 3.8点击Install。

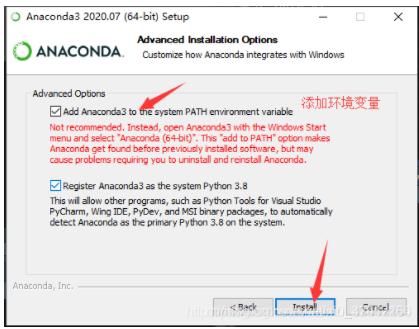


图7：安装Anaconda

如果忘记勾选了Add，添加环境变量，可以手动自己添加。配置环境变量：如果你在安装过程中勾选了Add PATH 那个选项，那么你不需要额外配置环境变量，但是如果你没有勾选（默认是不勾选的），那么你需要按照以下标准去配置（路径根据自己的安装位置变化）。

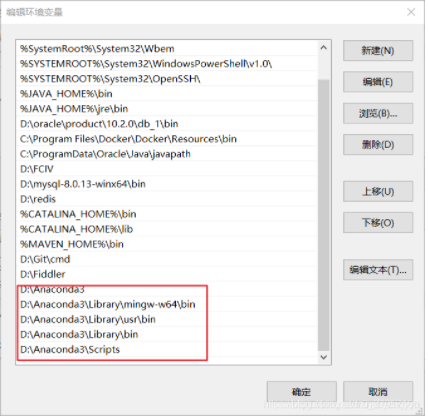


图8：添加环境变量

然后，它就开始自动安装了。

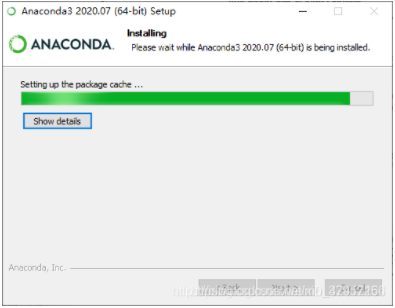


图9：安装Anaconda

绿色的进度条满了以后，点击Next。

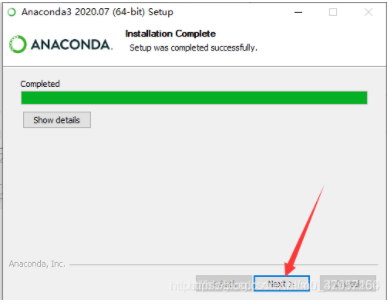


图10：安装Anaconda

还是点击Next。

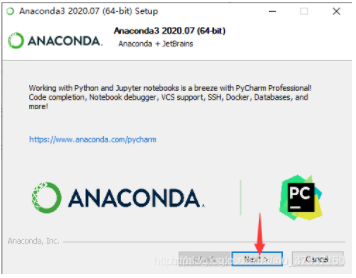


图11：安装Anaconda

这两个选项直接取消掉就行，跟安装一点关系没有，点击Finish。

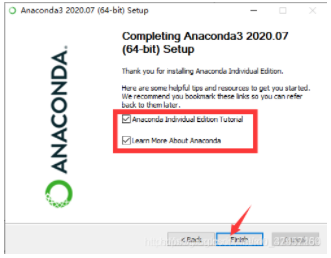


图12：完成安装Anaconda

## 验证是否成功安装

此时打开，开始->cmd，在终端中键入命令conda -V。进入cmd，进入路径，输入命令，查看是否安装成功。出现了如下效果，说明我们的安装已经完成，且Anaconda版本为4.9.2。



图13：验证安装Anaconda

或者也可以使用如下测试方法：测试安装的一种好方法是直接打开Jupyter Notebook。也可以通过Anaconda Prompt、Anaconda Navigator或Jupyter Notebook执行此操作。

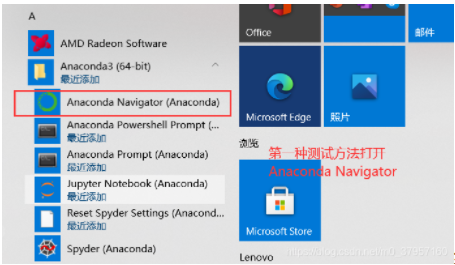


图14：验证安装Anaconda

1）通过Anaconda Navigator进行测试的结果如图15：

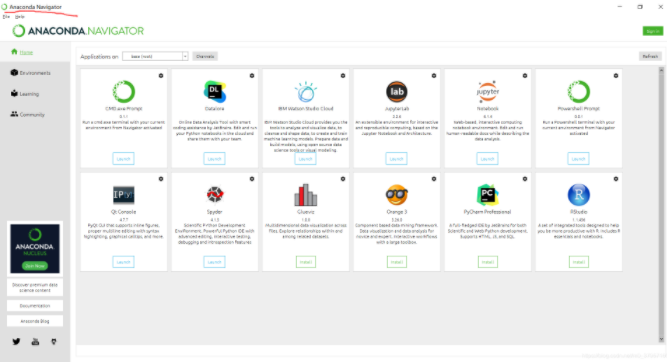


图15：验证安装Anaconda

2）通过Anaconda Prompt进行测试的方法如图16所示。

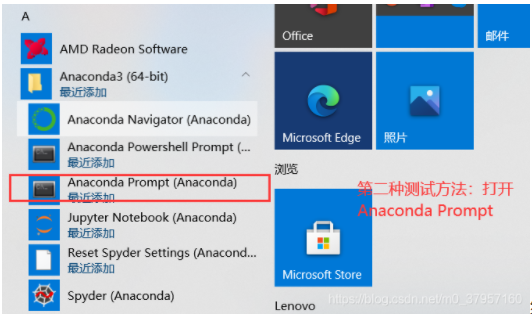


图16：验证安装Anaconda

结果如图17：

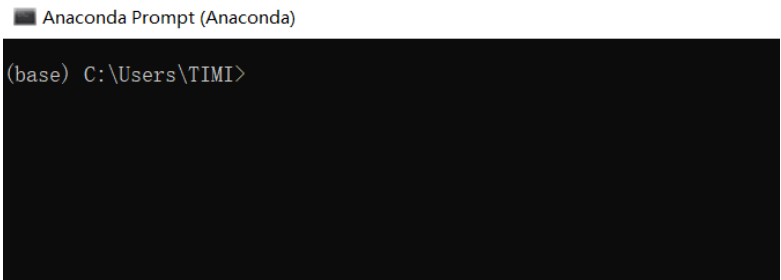


图17：验证安装Anaconda

3）也可以使用Jupyter Notebook，如图18。

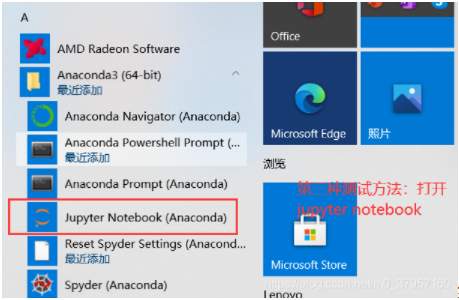


图18：验证安装Anaconda

结果如图19。

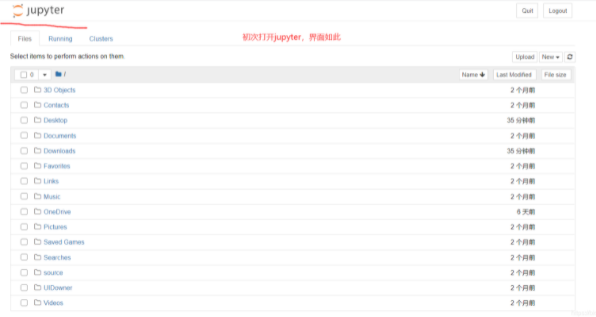


图19：验证安装Anaconda

## 使用Jupyter notebook

在cmd中直接输入命令打jupyter notebook。

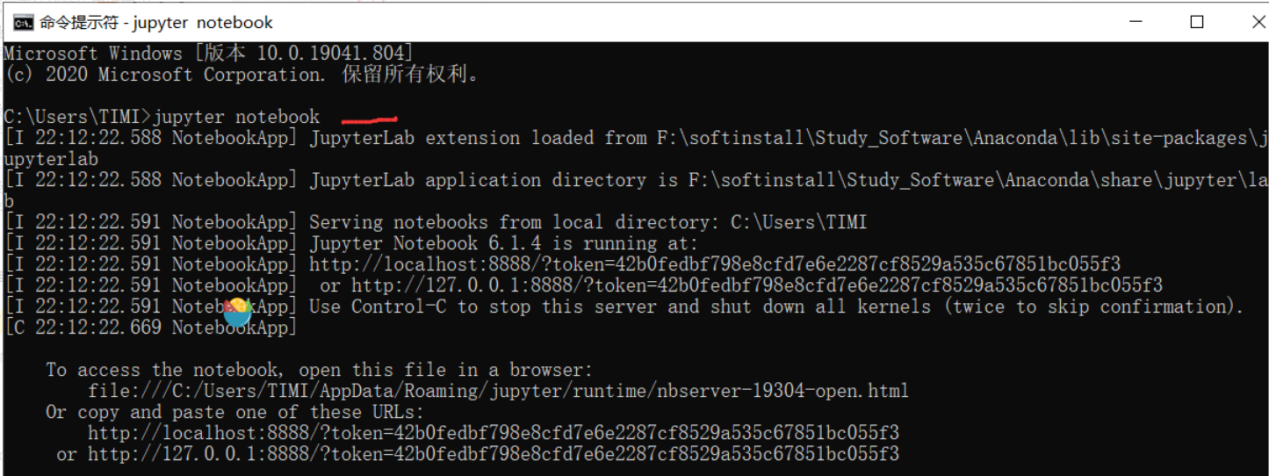


图20：打开Jupyter notebook

创建python代码编辑文本：new python3。

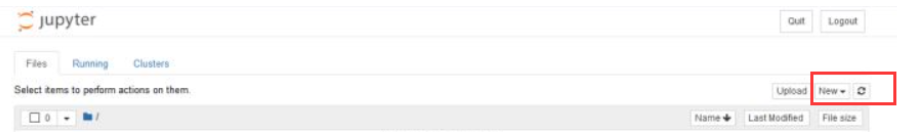


图21：新建Jupyter notebook

之后就是一个新界面如图22。

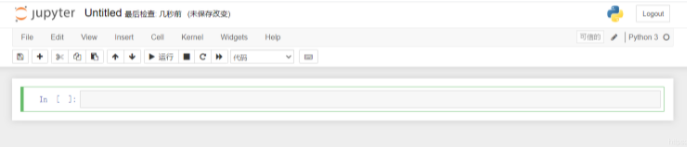


图22：Jupyter notebook代码界面

编写代码–运行，结果可直接显示在下方。

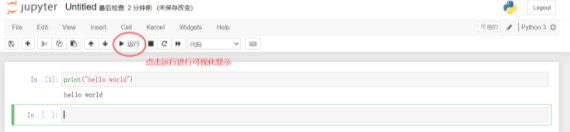


图23：Jupyter notebook执行代码

可以创建多个工作区、各工作区互不影响。



图24：Jupyter notebook代码工作区

# 实验一、使用基本数据类型

## 实验目的

熟悉Python开发环境；掌握Python语言三种数字类型及字符串类型的使用及其操作，能编写简单的Python程序。

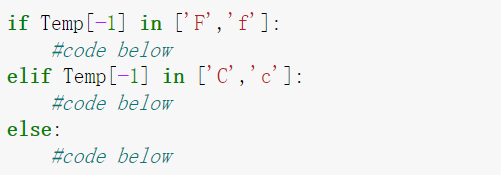
## 实验内容

**（1）编写Python程序，进行摄氏度和华氏度之间的温度转换。**

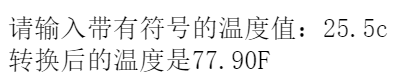
提示用户输入一个温度值。如输入一个末尾带有F（或f）的浮点数，则视之为华氏温度数，将其转化为末尾带有C的摄氏温度数；如输入一个末尾带有C（或c）的浮点数，则视之为摄氏温度数，将其转化为末尾带有F的华氏温度数；否则提示输入错误。

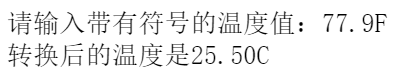
**提示：**

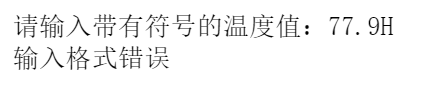
* 使用input函数提示用户输入一个温度值，使用format函数格式化输出
* 摄氏度和华氏度转换公式F = 1.8\*C+32（F华氏度，C摄氏度）
* 摄氏度输入格式为float+’c’/’C’，华氏度输入格式为float+’f’/’F’，其他输入提示格式错误如：



* 输入输出样例：







* 输入代码示例：



* 输出代码示例：



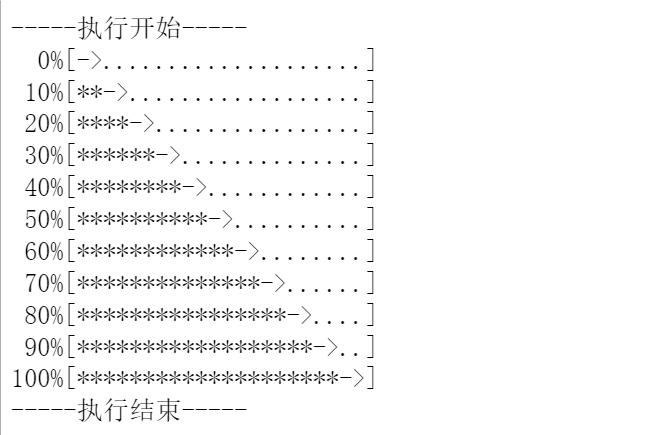




**（2）利用print()函数实现简单的非刷新文本进度条及单行动态刷新文本进度条。**

**提示：**

* 输出样例



* 引入Time库，使用time.sleep(0.1)可以使程序停止0.1秒后再继续运行
* 使用一个for循环，每次循环中，获取相应的百分比、星号长度、点号长度，依次打印百分比、左方括号、星号、—>、点号和右括号。如下所示，c表示百分比，a表示1个或多个星号，b表示1个或多个点号



## 实验报告要求

给出2个程序的程序代码。

# 实验二、程序控制结构

## 实验目的

理解程序设计的三种基本结构；掌握Python语言单分支结构、二分支结构以及多分支结构语句格式；熟练掌握for语句遍历循环以及while语句无限循环结构；掌握循环保留字：break和continue的使用；能编写简单的控制结构程序。

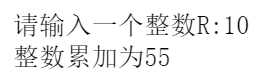
## 实验内容

**（1）整数累加：求1-R的和。**

输入一个正整数R，计算从1累加到R的和，并输出结果。

**提示：**

* 输入输出样例



**（2）编写一个“猜数字游戏”的程序**

在1到1000之间随机生成一个数，然后请用户循环猜测这个数字，对于每个答案只回答“猜大了”或“猜小了”，直到猜测准确为止，输出用户的猜测次数。

**提示：**

* 使用random库中的random.randint(1,1000)可以在1-1000中随机一个目标值（import random）

## 实验报告要求

给出“（1）整数累加”的代码和“（2）猜数字游戏”的竞猜过程截图。

# 实验三、股价可视化

## 实验目的

熟悉Python开发环境；掌握Pandas库和Matplotlib库的基本语法，能读取文件，并变换得到所需字段进行可视化。

## 实验内容

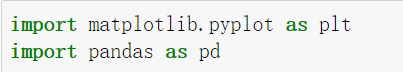
将可口可乐公司每日的收盘价以周、月为单位以平均值进行聚合，并将聚合结果以折线图、散点图、柱形图方式显示。

股价数据取自Coca\_data.csv，数据格式说明如下：

日期：DateTime（日期跨度2021.2.26 – 2022.2.25），收盘：Float，开盘：Float，高：Float，低：Float，交易量：Float(M)，涨跌幅:Float(%)

**提示：**

* 头文件：



* Pandas读取csv文件的方法：

data = pd.read\_csv(path\_csv)

* Pandas将日期列数据转化为datetime格式的方法：

data['日期'] = pd.to\_datetime(data['日期'])



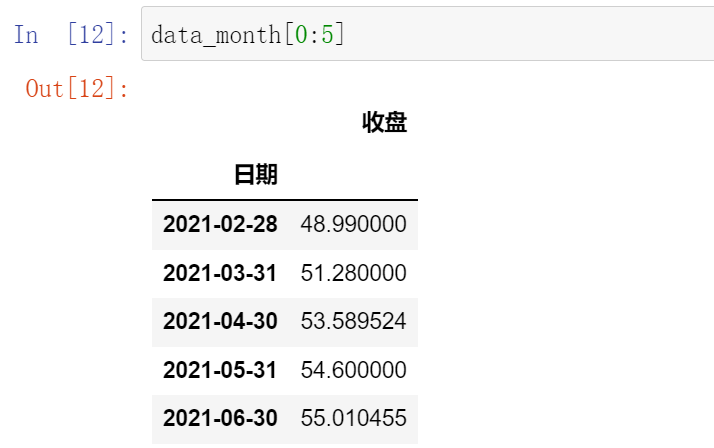
* 聚合函数

data\_month = data.groupby([pd.Grouper(freq='M', key='日期')]).agg({

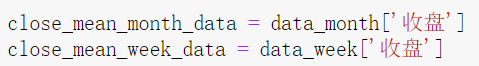
'收盘': 'mean'

})

将data根据日期以月作为频率进行分类，类内以收盘字段以均值进行聚合。按周聚合类似操作。



* pd.DataFrame转list的方法





* 绘图方法

折线图 plt.plot(x,y)，散点图 plt.plot(x, y, 'o')，柱形图 plt.bar(x,y)

## 实验报告要求

提供6张图，纵坐标为每日收盘价以周、月为单位的平均值，横坐标为周或月，显示形式为折线图、散点图和柱形图。

# 实验报告

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据科学导论实验一 | | | | |
| 题目 |  | 姓名 | 周联农 | 日期：2023/10/25 |
| 实验环境：python3.9 | | | | |
| 实验内容与完成情况： | | | | |
| 出现的问题： | | | | |
| 解决方案（列出遇到的问题和解决办法，列出没有解决的问题）： | | | | |
| 实验一结果： | | | | |
| 实验二结果： | | | | |
| 实验三结果： | | | | |