**项目一 Python数据分析环境初识**

**任务一 子任务2 Jupyter Notebook启动与诊断**

1.环境变量的作用？

2.在本地电脑中设置环境变量的流程？

**任务二 Jupyter Notebook编程与诊断**

使用Jupyter Notebook 输出使用的pandas版本号，并输出“Hello World”。

**项目二 Python科学计算库训练**

**任务一 初识NumPy**

1.（单选）NumPy包的核心是（B）

A. numpy

B. ndarray

C. 列表

D. 以上都不是

2.（判断）引入NumPy包可以使用import numpy进行引入，更多情况下，为了简化代码编写，使用import numpy as np进行引入。（√）

3.（简答）孙权招募贤才时曾说“能用众力，则无敌于天下矣；能用众智，则无畏于圣人矣”，这句话出自《三国志•吴书》，类比ndarray数组，请思考数组和普通变量相比有什么优势？

答案：（1）按照索引查询元素速度快

（2）能存储大量数据

（3）按照索引遍历数组方便

NumPy数据结构的优势：

（1）内存大小—Numpy数据结构占用的内存更小

（2）性能—Numpy底层是用C语言实现的，比列表更快

（3）运算方法—内置优化了代数运算等方法

（参考：https://zhuanlan.zhihu.com/p/469301475）

**任务二 创建NumPy数组**

1.（判断）NumPy中的array函数可以将指定的Python内置序列转换为数组，Python内置序列可以是列表、嵌套列表、元组、嵌套元组等指定的数据结构。（√）

2.（单选）下列哪个内置函数可以创建全0数组？（C）

A. arange

B. ones

C. zeros

D. empty

3.（多选）下列哪个描述表示浮点型？（AB）

A. float64

B. np.float64

C. int32

D. string\_

4.（单选）下列哪个属性可以获得数组的维度？（A）

A. ndarray.ndim

B. ndarray.shape

C. ndarray.dtype

D. ndarray.size

**任务三 数组的基本操作**

1.（单选）用于求绝对值的函数是（B）

A. np.power

B. np.abs

C. np.divide

D. np.add

2.（判断）ndarray对象的内容可以通过索引来访问和修改，可以基于 0 - n 的下标进行索引。（√）

3.（简答）总结冒号的使用方法。

答案：

* 如果在中括号中只放置一个参数，如[2]，将返回与该索引相对应的单个元素。
* 如果为[2:]，表示从该索引开始以后的所有项都将被提取。
* 如果使用了两个参数，如[2:7]，那么则提取两个索引(不包括终止索引)之间的项。
* 如果中括号中出现了逗号，可以将逗号前面的操作理解为对第0轴的操作，逗号后面的操作理解为对第1轴的操作
* 如果只出现一个冒号，表示选择全部元素（arr[:]中冒号表示选择arr全部元素，arr[:,1]中冒号表示选择第0轴全部元素，arr[1,:]中冒号表示选择第1轴全部元素）

**项目三 Python数据分析库训练**

**任务一 pandas数据结构**

1.（判断）当DataFrame数据中出现各列数据类型不同时，转换为二维数组后返回的数组不会兼顾所有的数据类型。（×）

**任务二 pandas索引操作**

1.（单选）设置索引使用哪种方法（D）。

A. merge()方法 B. concat()方法 C. to\_datetime()方法 D.set\_index()方法

2.（判断）Pandas中可以通过行索引或行索引位置的切片形式选取行数据。（√）

3.（判断）两个索引不一致的Series进行算数运算会出错。（×）

**任务三 pandas数据运算**

1.（选择）describe方法中的25%表示什么？（D）

A. 行数 B. 标准差 C. 最大值 D. 第一四分位数

**任务四 层次化索引**

1.（判断）只有行方向上可以进行层次化索引。（×）

**任务五 pandas可视化**

1.（填空）线形图由（plot()）函数来绘制。

2.（填空）柱状图由（plot(kind=’bar’)）函数来绘制。

3.（填空）直方图由（hist()）函数来绘制。

4.（填空）密度图由（plot(kind=’kde’)）函数来绘制。

5.（填空）散点图由（plot(kind=’scatter’)）函数来绘制。

**项目四 Python大数据分析基础综合应用**

**任务一 需求分析**

1.（多选）需求分析包括三个阶段（ABC）

A. 发现问题

B. 需求确认

C. 需求处理

D. 需求回顾