

2.5 作业

2023 年 9 月 18 日

作业

1. 定义变量 (1 分)

- 给一个变量名为“year”的变量赋值字符串: ‘2023’
- 给一个变量名为“month”的变量赋值字符串: ‘09’
- 定义一个变量表示日期, 赋值字符串: ‘01’
- 拼接上述三个字符串, 并使用‘/’来连接

[]:

2. 创建一个列表 (1 分)

- 请创建一个列表变量, 名字叫 college, 包含以下元素: “外语学院”, “国际交流学院”, “教育学院”, “传媒学院”
- 使用 print 语句输出上述列表
- 使用 append 语句在该列表后追加一个新的元素: “商学院”

提示: 在调试代码过程中, 如果出现语句错误, 如“SyntaxError: invalid character in identifier”, 讲中文逗号, “改成英文逗号”。

[]:

3. 创建一个字典 (1 分)

- 使用字典创建一个变量, 变量名为 account_info, 包含账户名: “user_01”, 和密码“123456”, 然后打印变量。
- 给上面这个字典变量, 添加一个新的键-值对, 键为“user_02”, 值为“654321”。

[]:

4. 计算课程总成绩（1 分）

相信很多老师在授课的第一节课都会告诉学生，本学期的考核和成绩评定方式，例如：

- 课程总成绩 = 平时成绩（30%）+ 测试（70%）
- 平时成绩 = 出勤（10%）+ 课内作业（20%）
- 测试 = 期中实验报告（10%）+ 期末实验报告（60%）

现在学校要求我们编写一个程序，让学生输入他本学期的 16 周的出勤次数、平时作业成绩、期中测验分数和期末卷面成绩来计算出他的总成绩。请告知一下同学他的分数：

```
student_01 = {'name': 'zhangxiaoxian', # 姓名
              'attendance_numbers': 14, # 出勤数
              'homework_score': 75, # 平时作业分数
              'mid_term_exam_score': 80, # 期中考试分数
              'final_term_score': 85} # 期末考试分数

student_02 = {'name': 'wangpeng',
              'attendance_numbers': 3,
              'homework_score': 80,
              'mid_term_exam_score': 90,
              'final_term_score': 95}

student_03 = {'name': 'lifeng',
              'attendance_numbers': 12,
              'homework_score': 60,
              'mid_term_exam_score': 70,
              'final_term_score': 75}
```

[]:

5. 循环打印（1 分）

请利用循环 for 语句和 print 语句，依次对 list 中的每个名字打印出 Hello, xxx!:

```
L = ['Bart', 'Lisa', 'Adam']
```

[]:

6. 计算平均值 (1 分)

均值是描述性统计方法中最常见的指标，它主要用来计算全体样本或抽样的平均值。

$$\mu = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

试用变量、列表 list 和循环 for 语句计算数列 [1, 2, 3, 4, 5] 的平均值。

[]:

7. 计算方差 (1 分)

参照上述平均值的计算方法，进一步计算数列 [1, 2, 3, 4, 5] 的方差。方差公式为：

$$\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \cdots + (x_n - \mu)^2}{n}$$

提示：使用变量、列表 list 和循环 for 语句，也可使用函数 def 语句

[]:

8. 计算最大值 (1 分)

试用变量、循环 for 语句和判断 if 语句计算数列 [1, 5, 3, 4, 2] 中的最大值。

[]:

9. 设计一个计算均方误差的函数 (1 分)

在回归分析中，均方误差 (MSE) 是一个常用的指标。

它计算的是预测数据和原始数据对应点误差的平方和的均值，它主要用来评估模型的预测能力。

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

这里 \hat{y}_i 是指模型预测的数据序列， y_i 是指原始数据序列， \sum 是累加符号。

设计一个函数名为 `mse`，输入参数为 `original_y` 和 `predicted_y`，返回值为计算的结果。

请问，当 `original_y=[10.1, 9.8, 10.5, 10.0, 10.3]`, `predicted_y=[10, 10, 10, 10, 10]` 时，MSE 的值是多少？

提示：使用变量、for 循环语句、def 语句和 return 语句来设计函数，传入变量是两个列表，然后调用并测试函数是否成功运行

[]:

10. 设计一个返回逆序列表的函数 (1 分)

请设计一个函数，函数名字为 `reverse`，它的输入变量为一个列表，返回这个列表的逆序，例如

```
>>> l = [1, 2, 3, 4, 5]
>>> reverse(l)
[5, 4, 3, 2, 1]
```

[]:

加分题：整数编码 (1 分 * 2 = 2 分)

在数据挖掘模型中，数据集的属性如果是字符型，需要将其变换为数字，最常用的方法叫整数编码。

它有两个步骤：1. 建立一个字符串和整数一一对应的字典。2. 使用循环，遍历所有字符串，将字符串替换为整数。

现有有一个西瓜数据集，请按照下面规则对其进行整数编码。

属性：

色泽 1-3代表 浅白 青绿 乌黑

根蒂 1-3代表 稍蜷 蜷缩 硬挺

敲声 1-3代表 清脆 浊响 沉闷

纹理 1-3代表 清晰 稍糊 模糊

脐部 1-3代表 平坦 稍凹 凹陷

触感 1-2代表 硬滑 软粘

标签：

好瓜 1代表 是 0 代表 不是

任务 1:

`w = ['青绿', '蜷缩', '浊响', '清晰', '凹陷', '硬滑', '0.697', '0.460', '是']`

列表各元素所对应的属性为‘色泽’，‘根蒂’，‘敲声’，‘纹理’，‘脐部’，‘触感’，‘密度’，‘含糖率’，‘好瓜’ 1. 建立各个属性的字符为键，整数为值的字典 2. 将这个列表的各个元素从字符串转换成数值：[2, 2, 2, 1, 3, 1, 0.697, 0.460, 1]

[]:

任务 2:

如下是一个嵌套的列表，嵌套的第一个列表为属性和标签，每一个列表的第一个元素为序号。请只转换需要转换的字符串为数值，其他不变。

```
w02 = [['编号', '色泽', '根蒂', '敲声', '纹理', '脐部', '触感', '密度', '含糖率', '好瓜'],
        ['1', '青绿', '蜷缩', '浊响', '清晰', '凹陷', '硬滑', '0.697', '0.460', '是'],
        ['2', '乌黑', '蜷缩', '沉闷', '清晰', '凹陷', '硬滑', '0.774', '0.376', '是'],
        ['3', '乌黑', '蜷缩', '浊响', '清晰', '凹陷', '硬滑', '0.634', '0.264', '是'],
        ['4', '青绿', '蜷缩', '沉闷', '清晰', '凹陷', '硬滑', '0.608', '0.318', '是'],
        ['5', '浅白', '蜷缩', '浊响', '清晰', '凹陷', '硬滑', '0.556', '0.215', '是'],
        ['6', '青绿', '稍蜷', '浊响', '清晰', '稍凹', '软粘', '0.403', '0.237', '是'],
        ['7', '乌黑', '稍蜷', '浊响', '稍糊', '稍凹', '软粘', '0.481', '0.149', '是'],
        ['8', '乌黑', '稍蜷', '浊响', '清晰', '稍凹', '硬滑', '0.437', '0.211', '是'],
        ['9', '乌黑', '稍蜷', '沉闷', '稍糊', '稍凹', '硬滑', '0.666', '0.091', '否'],
        ['10', '青绿', '硬挺', '清脆', '清晰', '平坦', '软粘', '0.243', '0.267', '否'],
        ['11', '浅白', '硬挺', '清脆', '模糊', '平坦', '硬滑', '0.245', '0.057', '否'],
        ['12', '浅白', '蜷缩', '浊响', '模糊', '平坦', '软粘', '0.343', '0.099', '否'],
```

```
['13', '青绿', '稍蜷', '浊响', '稍糊', '凹陷', '硬滑', '0.639', '0.161',  
↪ '否'],  
['14', '浅白', '稍蜷', '沉闷', '稍糊', '凹陷', '硬滑', '0.657', '0.198',  
↪ '否'],  
['15', '乌黑', '稍蜷', '浊响', '清晰', '稍凹', '软粘', '0.360', '0.370',  
↪ '否'],  
['16', '浅白', '蜷缩', '浊响', '模糊', '平坦', '硬滑', '0.593', '0.042',  
↪ '否'],  
['17', '青绿', '蜷缩', '沉闷', '稍糊', '稍凹', '硬滑', '0.719', '0.103',  
↪ '否']]
```

[]: