**上海市“星光计划”**

**第十届职业院校职业技能大赛**

“Python程序设计”项目

**竞赛样题**

上海市星光计划组委会竞赛办公室

二〇二二年十一月

**竞赛任务及要求**

# 模块A：Python编程基础

## 变量赋值

### 任务1

1. 参赛选手收到一条信息message，内容为“Welcome to the python match!”，请打印出对应信息的数据类型。
2. 定义变量message，将信息传递给它，将其数据类型打印展示。
3. 代码保存地址为：/root/python1/var.py

### 任务2

1. 参赛前，组委会需要对选手进行信息确认，要求学生输入对应姓名和学号，即可得到输出相关信息，实例如下：

终端输入命令：python/root/python1/student.py Mike 20210420

终端输出结果：请确认您的信息！姓名：Mike，学号：20210420

**选手须知：**

1. 使用argv[]为变量name、num进行外部传参；
2. 使用format进行格式化输出；
3. 代码保存地址为：/root/python1/student.py；

## 运算符使用

### 任务1

1. 现有某学生语文、数学、外语三门成绩分别为99、80、76，要求使用变量及运算符知识，将语文成绩赋值给a，将数学成绩赋值给b，将外语成绩赋值给c，计算学生平均成绩并输出；

1）本题中成绩按照整型进行运算，结果按照四舍五入进行取整；

2）代码保存地址为：/root/python1/average.py；

### 任务2

1. 根据提供的某班级数学成绩单/root/python\_data/math.csv，计算并打印班级平均成绩。

math.csv数据：80, 89, 90, 45, 66, 87, 40, 78, 65, 100, 96, 86, 78, 90, 68, 90

1）数据中只有一个字段为：数学成绩，数据类型为int;

2）使用with open方式读取文件数据；

3）要求输出结果使用int()向下取整；

4）代码地址保存为：/root/python1/score.py

## 大小写转换

### 任务1

1. 学校英语角的标题为“Congratulations on winning the first prize!”，为了强调内容，现在需要将标题更改为全部大写，编写程序使用内置函数将标题进行全部大写转换，运行即可输出结果。
2. 代码保存地址为：/root/python1/upswitch.py

### 任务2

1. 编写一个字符串大小写转换程序，对于接收的字符串中各字母，进行大小写互换；对于其他字符则直接输出，不进行转换，实例如下：

终端输入：python/root/python1/switch.py ABcdeF/

终端输出：abCDEf/

1）使用外部传参的方式进行字符串输入；

2）要求使用ASCII码转换字符串大小写，ASCII是基于拉丁字母的一套电脑编码系统，例如a对应ASCII码为97，A对应ASCII码为65；

3）禁止使用lower()、upper()等内置函数；

4）代码保存地址为：/root/python1/switch.py;

## 进制转换

计算机中IP地址是32位二进制数字组成的数字序列，常用的表示方法是将二进制数字从末尾往前每4个二进制数字作为一个字节转换成10进制数字进行表示，因此平时常见的IP表示为12个阿拉伯数字，每3个数字为一组，组与组之间用英文句点“.”来分隔，例如192.168.12.144。

### 任务1

1. 编写代码，对以上十进制数值“192”，进行二进制转换并打印转换结果；

1）使用内置函数对其进行转换，保留前缀；

2）代码保存地址为：/root/python1/conversion.py

### 任务2

1. 编程进制转换程序代码，要求可以对输入的IP地址进行二进制转换并输出结果，实例如下：

终端输入：python/root/python1/ip.py 192.168.12.144

终端输出：11000000101010000000110010010000

1）使用外部传参的方式为变量ip进行赋值；

2）要求IP地址中每组数值转换后长度为8，长度不足部分左侧补0,推荐使用内置函数补齐字符串;

3）输出结果忽略二进制前缀，连接每组数据，最后输出结果长度为32位；

4）代码保存地址为：/root/python1/ip.py

## 运算符使用

### 任务1

1. 现有直径50cm的下水道井盖，使用python运算知识，求其面积并打印结果。
2. 其中π直接引用数学库中的pi，pi即为math中的常量；
3. 输出结果中不包含单位，结果为整型；
4. 代码保存地址为：/root/python1/well.py

### 任务2

1. 已知上网流量1GB=1024MB，19元包月套餐中通用流量为15G，若超出流量按照0.29元/MB计，不足1M的按照1M算，本月流量已超4.25G，求本月总账单定打印结果。

1）结果保留两位小数（单位元忽略）。

2）代码保存地址为：/root/python1/flow.py

### 任务3

1. 根据上题中的流量计费标准，编写函数，要求输入对应流量数值，即可得到对应账单，并输出账单结果。

1）定义一个函数bill()进行费用计算，参数为n；

2）使用外部传参方式对n（单位GB）进行赋值，数值类型为float;

3）使用内置函数对账单结果进行向下取整，注意流量不足1M的按照1M计算；

4）代码保存地址为：/root/python1/bill.py

### 任务4

1. 编写程序满足如下需求：输入一个整数n,计算该整数的12次方，并输出结果。

1）要求调用math模块函数进行次方计算；

2）使用sys.argv方式传递外部参数;

3）结果取整；

3）代码保存地址为：/root/python1/power.py

### 任务5

1. 已知三角形的两边长a，b为5和6，两边夹角C为30度，编写程序计算第三边长并输出。

1）要求调用python内置函数进行平方根计算；

2）三角形边角关系为：$a^{2 }=b^{2}+c^{2}-2bc\*\cos(A)$；

3）Python中cos(x)函数中x是一个数值，返回的是x的弧度的余弦值；

4）弧度与角度关系为：弧度 = 角度\*PI/180；角度 = 弧度\*180/PI;

5)使用format格式进行字符串格式化输出，保留两位有效小数；

6）代码保存地址为：/root/python1/trangle.py

## 控制语句

### 任务1

1. 编写程序，输入整数，判断对应整数年份是否为闰年并输出结果，实例如下：

终端输入：python/root/python1/year.py 2018

终端输出：2018年不是闰年

终端输入：python/root/python1/year.py 2020

终端输出：2020年是闰年

1）闰年判定方法：被4整除，但不能被100整除；或者能被400整除；

2）使用外部传参的方式将整数传递给变量year；

3）使用format格式进行字符串格式化输出；

4）代码保存地址为：/root/python1/year.py

### 任务2

1. 现有成绩分级，成绩小于60分为不及格，大于80分为优秀，其他为良好，要求根据学生成绩变量score为78分，打印成绩等级。

1）使用变量score进行分数赋值；

2）使用内嵌if判断进行程序编写；

3）成绩等级为“不及格”、“良好”、“优秀”；

4）代码保存地址为：/root/python1/grade.py

### 任务3

1. 基于上一题规则，编写成绩分级程序，根据输入的成绩数值，进行自动分级，并打印成绩等级。实例如下：

终端输入：python/root/python1/autograde.py 111

终端输出：无效分值

1）使用外部传参(sys)的方式将分数传递给变量score；

2）使用内嵌if判断进行程序编写；

3）成绩等级为“不及格”、“良好”、“优秀”；

4）成绩超过100或小于0分时，输出结果为“无效分值”。

5）代码保存地址为：/root/python1/autograde.py。

### 任务4

1. 编写程序，使用while循环计算0-100所有的奇数之和sum，打印结果；

1)创建变量sum，初始值为0；

2)代码保存地址为：/root/python1/sum.py

### 任务5

1. 编程求和程序，要求根据输入的参数闭间值a,b（闭区间）,计算两者之间的所有奇数之和sum。实例如下：

终端输入：python/root/python1/region.py 7 9

终端输出：16

1）使用外部传参方式（sys）进行区间值输入;

2）传递的第一个参数为a，第二个参数为b，要求判断a小于等于b，否则输出“参数区间错误”;

3）代码保存地址为：/root/python1/region.py

# 模块B：Python综合编程

## 用python语言实现九九乘法表

**项目背景及要求**

学习编程，循环是必备能力之一，本题需要使用while循环结构编写程序打印九九乘法表（正三角），请你按程序文件的注释，将Python代码补充完整，运行程序，输出的结果。输出的结果形式如下：

1）代码保存地址为：/root/python2/form.py，补全代码；

2）使用占位符%对十进制整数进行结果格式化输出，要求整型占位符宽度为2，默认右对齐方式输出，使用end关键字在每个计算结尾处添加空格；

3）满足循环条件，则执行程序，更新变量循环参数，否则结束循环；

请你按程序文件的注释，将Python代码补充完整，运行程序，输出的结果。输出的结果形式如下图：

1\*1= 1

1\*2= 2 2\*2= 4

1\*3= 3 2\*3= 6 3\*3= 9

1\*4= 4 2\*4= 8 3\*4=12 4\*4=16

1\*5= 5 2\*5=10 3\*5=15 4\*5=20 5\*5=25

1\*6= 6 2\*6=12 3\*6=18 4\*6=24 5\*6=30 6\*6=36

1\*7= 7 2\*7=14 3\*7=21 4\*7=28 5\*7=35 6\*7=42 7\*7=49

1\*8= 8 2\*8=16 3\*8=24 4\*8=32 5\*8=40 6\*8=48 7\*8=56 8\*8=64

1\*9= 9 2\*9=18 3\*9=27 4\*9=36 5\*9=45 6\*9=54 7\*9=63 8\*9=72 9\*9=81

部分代码如下：

# -\*- coding:utf-8 -\*-

i=1

while ?:

j?

while ?:

print(?,end=' ')

?

print("")

?

代码如下：

# -\*- coding:utf-8 -\*-

i=1

while i<10:

j=1

while j<=i:

print("%d\*%d=%2d"%(j,i,j\*i),end=' ')

j+=1

print("")

i+=1

# 检测条件：

# 1.检测文件中，%d\*%d=%2d格式化输出命令是否补全、end关键字是否存在；

# 2.运行程序，检测程序是否可以成功运行；

# 3.最后检测结果，3\*5=15、9\*9=81（此部分检测以实际内容为准），以上条件均正确，是本题正确。

## 手机号运营商分类

**项目背景**

某学校要统计所有学生手机号为哪些运营商的情况，手工统计非常耗时。请你按程序文件的注释，将Python代码补充完整，使用程序完成统计任务，运行程序并输出结果。

**选手须知**

了解三大运营商的号段分布；

1）中国联通：130，131，132，155，156，185，186，145，176

2）中国移动：134, 135 , 136, 137, 138, 139, 147, 150, 151,152, 157, 158, 159, 178, 182, 183, 184, 187, 188

3）中国电信：133,153,189

### 任务1

1. 判断手机号是否合法，依次按照长度(11位)、字符(只包含数字)、号段进行判断归类，若一个手机号长度和字符都不合法，按照顺序，需要将其划分到长度不合法里面；

### 任务2

1. 编写代码，判断手机号是否重复；

### 任务3

1. 编写代码，根据给定的学生信息进行手机号数据文件导入；

### 任务4

1. 运行输出为：移动数量为？个,联通数量为？个,电信数量为？个,长度不合法数量为？个,号段不合法数量为？个,字符不合法为？个。

### 任务5

1. 代码保存地址为：/root/python2/phone.py

### 任务6

1. 手机号文件地址为：/root/python\_data/phone.txt

手机文件phone.txt：

183103968989

1802123456478

15602340044

15608590646

15201260001

13912345678

139123456785

1391234a678

23912345678

19012345678

## 人民币大小写切换

**项目背景**

工作中常常需要报销差旅费等其他费用，但是由于报销单需要进行大小写规范，因此很多人进行大写数字书写时，常常需要进行资料查找。这里需要选手设计一个万位以内的数值大写转换（例如9999.99，最高为千位，最多两位小数），请你编写程序文件cnyswitch.py，运行程序，输入小写数值，返回对应的大写数字。实例如下：

终端输入：python/root/python2/cnyswitch.py 1234.56

终端输出：壹仟贰佰叁拾肆圆伍角陆分

**选手须知**

1.重点考核选手根据要求，自定义代码规则，创建函数满足对应功能；

2.中文金额数字大写分别为“[壹](https://baike.baidu.com/item/%E5%A3%B9/37883%22%20%5Ct%20%22/Users/zhangdabao/Documents%5C%5Cx/_blank)、贰、叁、肆、伍、陆、柒、捌、玖、拾、佰、仟、圆、角、分、零、整”。

3.按照数位顺序表，万以内的数据值转换，最高位为千位，数值最多两位小数，即分位。

4.代码保存地址：/root/python2/cnyswitch.py

## 查找关系密切的组合

**项目背景**

作为一个电影爱好者，常常会对电影相关信息进行分析，我们会发现很多演员会在不同的电影里进行合作，如果想要找出合作最为亲密的两个演员，手工查找分析是非常慢的，请你按程序文件的注释，将Python代码补充完整，从/root/python\_data/film.xlsx中找出合作密切的两位演员，并用程序输出结果；

**选手须知**

要了解频繁项集：经常一起出现的物品的集合。如果某个项集是频繁的，那么它的所有子集都是频繁的；如果某个项集不是频繁的，那么它的所有超集都不是频繁的。这一点是避免项集数量过多的重要基础，使得快速计算频繁项集成为可能。

1)代码文件地址为：/root/python2/film.py

2)数据文件地址为：/root/python\_data/film.xlsx（注意文件是否正确,xlsx文件借助vscode上Excel Viewer进行查看，选中文件右键’open preview‘即可）

3)运行代码，结果输出实例如下：

终端输入：python/root/python2/film.py

终端输出：关系最好的2个演员是演员a,演员b,他们共同出演的电影数量是c.

备注：其中a,b,c为计算后的真实内容，这里仅做结果样式介绍。

|  |  |  |
| --- | --- | --- |
| 电影名称 | 导演 | 演员 |
| 电影1 | 导演1 | 演员1，演员2，演员3，演员4 |
| 电影2 | 导演2 | 演员3，演员2，演员4，演员5 |
| 电影3 | 导演3 | 演员1，演员5，演员3，演员6 |
| 电影4 | 导演1 | 演员1，演员4，演员3，演员7 |
| 电影5 | 导演2 | 演员1，演员2，演员3，演员8 |
| 电影6 | 导演3 | 演员5，演员7，演员3，演员9 |
| 电影7 | 导演4 | 演员1，演员4，演员6，演员7 |
| 电影8 | 导演1 | 演员1，演员4，演员3，演员8 |
| 电影9 | 导演2 | 演员5，演员4，演员3，演员9 |
| 电影10 | 导演3 | 演员1，演员4，演员5，演员10 |
| 电影11 | 导演1 | 演员1，演员4，演员3，演员11 |
| 电影12 | 导演2 | 演员7，演员4，演员9，演员12 |
| 电影13 | 导演3 | 演员1，演员7，演员3，演员13 |
| 电影14 | 导演4 | 演员10，演员4，演员9，演员14 |
| 电影15 | 导演5 | 演员1，演员8，演员11，演员15 |
| 电影16 | 导演6 | 演员14，演员4，演员13，演员16 |
| 电影17 | 导演7 | 演员3，演员4，演员9 |
| 电影18 | 导演8 | 演员3，演员4，演员10 |

代码如下：

# 运行结果：关系最好的两个演员是('演员？','演员？'),他们共同出演的电影数量是？。

## 根据身份证号计算退休时间

**题目背景**

已知当前法定退休年龄为男性60周岁，女性55周岁，现要求根据某公司员工身份证信息进行年龄和退休时间统计，请你按程序文件的注释，将Python代码补充完整，实现年龄和退体时间的计算，运行代码，保存运行结果。结果中单条数据实例如下（字段为：姓名，身份证号，当前年龄，退休日期）：

**赵峰平 571115196702015183 54 2022-02-01**

**选手须知**

1）注意数据文件的读取。

2）日期时间对象和字符串之间的相互转换。

3）获取当前时间可以使用datetime.now()

4）字符串格式化与排版。

5）员工信息数据/root/python\_data/staff.csv，数据字段为姓名、身份证号；

6）身份证号码是18位数，其中第7-10位是出生年份，第11-12位是出生月份，第13-14位是出生日，第17位是性别判断，奇数为男，偶数为女。

7）代码保存地址：/root/python2/retire.py

8）结果保存地址：/root/python\_data/staffretire.txt，字段为姓名，身份证号，当前年龄，退休日期；

员工信息staff.csv

姓名,身份证

赵峰平,571115196702015183

吴全国,485246196606014798

钱峰华,093316200603011668

李国,951901200706017489

钱天,730403198206013216

王磊,095634199910016091

周延天,434243196111018578

孙平,854913200009011436

郑明天,011999199509017300

王延天,210592199007014318

吴京华,928008198804016098

王平,220712200405015323

李峰华,657717199107011668

钱雨平,207434197601016147

吴华,949744198603011057

孙天,745368196906013182

周华,687449197507016772

李天,258883198608018052

钱华,478673197512018452

郑明平,315247198201018651

# 模块C：Python爬虫程序设计

## 论坛数据采集

**项目背景**

小王接到公司领导一个任务，要求小王把某论坛内所有的帖子，用户名，积分，等级，标题等内容全部复制下来，为数据分析部门做好数据支撑，如果手工复制将是一个巨大的工作量，请你帮用python语言，帮小王编写一个自动的数据爬取程序，将相关内容自动采集下，并生成一个data.txt文件保存。

**要求如下：**

要求代码书写规范，每行代码均有注释说明，采集的结果正确；



# 模块D：Python数据可视化

模块中评价分（Judgement）的打分方式：本模块由3名裁判为一组，各自单独评分，计算出平均权重分，除以3后再乘以该子项的分值计算出实际得分，主要评分依据包括程序能否正常运行、可视化结果是否美观合理、是否符合行业标准、代码注释是否全面、代码编写是否规范等。

## 销售季度销售额展示

**项目背景**

绘制折线图表，以季度销售额为例，1-3月份销售额为257,4-6月份销售额为301,7-9月份销售额为428,10-12月份销售额为475，参考样图如下：



**要求如下：**

要求代码书写规范，每行代码均有注释说明，按照指定正确展示；

## 各省大米销量的占比分析

**项目背景**

通过饼形图分析2020年11月各省大米销量的占比情况

**要求如下：**

创建绘图对象:设置画布的宽度和高度为7×5英寸

绘制饼形图:省份作为每一块饼形图外侧显示的说明文字;销量作为每一块饼形图的数据

设置图表标题:图表标题为'2020年11月各省大米销量占比情况分析'，字号18



**要求如下：**

要求代码书写规范，每行代码均有注释说明，按照指定正确展示；

## 散点图绘制

绘制简单散点图，填写下列代码，生成如下参考样图形式：

import numpy as np # 导入numpy模块

import matplotlib.pyplot as plt # 导入matplotlib中的pyplot模块

x = np.random.randint(???) # 随机生成8个[0,10)之间的整数

y = np.random.randint(???) # 随机生成8个[0,10)之间的整数

plt.scatter(x, y) # 绘制散点图

plt.show() # 显示图表



**要求如下：**

要求代码书写规范，每行代码均有注释说明，按照指定正确展示；

## 消费数据异常值处理

通过箱形图判断异常值，现有一份消费数据如下，将如下数据绘制成箱形图，并查找数据中的异常值，填写下列代码，生成如下参考样图形式：

import pandas as pd # 导入pandas模块

import matplotlib.pyplot as plt # 导入matplotlib中的pyplot模块

# 导入Excel文件

df = pd.read\_excel('./tips.xlsx')

# 手动添加黑体字体，解决中文乱码问题

plt.rcParams['font.sans-serif'] = ['？？？']

# 绘制箱形图

plt.boxplot(x=df['总消费'], # 指定绘制箱形图的数据

 ？=0.3, patch\_artist=True, # 箱子宽度0.3，填充箱体颜色

 # 显示均值，箱子的填充色为蓝绿色

 showmeans=True, boxprops={'facecolor': 'c'},

 # 指定异常值的填充色、边框色和大小

 flierprops={'markerfacecolor': 'r', 'markeredgecolor': 'r', 'markersize': 5},

 # 指定中位数的线条样式（虚线）、线宽和颜色

 medianprops={'linestyle': '？？', 'linewidth': 2, 'color': '？？'},

 # 指定均值点的标记符号（竖六边形）、填充色和大小

 meanprops={'marker': 'h', 'markerfacecolor': 'b', 'markersize': 8},

 labels=['']) # 去除x轴刻度值

plt.show() # 显示图表



## 二手房房价分析

现有二手房信息，要求读取数据，对数据进行清洗，并计算所有区域的二手房均价，并绘图展示，填写补充下列代码，生成如下参考样图形式：

import pandas as pd # 导入pandas模块

pd.set\_option('display.unicode.east\_asian\_width', True) # 解决数据输出时列名不对齐的问题

# 导入csv数据文件

house = pd.read\_csv('./house.csv')

print(house.head()) # 输出前5条数据



**要求如下：**

要求代码书写规范，每行代码均有注释说明，按照指定正确展示；