

湖南生物机电职业技术学院
学生专业技能考核题库

大数据技术专业

湖南生物机电职业技术学院
二〇二三年 九 月

目 录

一、专业基本技能	4
模块一 程序设计应用	4
项目 1: 小学生数学辅助学习系统	4
1. 试题编号: 1-1-1 《小学生数学辅助学习系统》关键算法	4
项目 2: 帮你算系统	6
2. 试题编号: 1-2-1 《帮你算系统》关键算法	6
项目 3: 网络选拔赛题库系统	7
3. 试题编号: 1-3-1 《网络选拔赛题库系统》关键算法	7
项目 4: 图形体积计算系统	9
4. 试题编号: 1-4-1 《图形体积计算系统》关键算法	9
项目 5: 中国结图案打印系统	11
5. 试题编号: 1-5-1 《中国结图案打印系统》关键算法	11
项目 6: 智能统计系统	13
6. 试题编号: 1-6-1 《智能统计系统》关键算法	13
项目 7: 儿童智力游戏	15
7. 试题编号: 1-7-1 《儿童智力游戏》关键算法	15
项目 8: 商品销售系统	16
8. 试题编号: 1-8-1 《商品销售系统》关键算法	16
项目 9: 幼师培训系统	18
9. 试题编号: 1-9-1 《幼师培训系统》关键算法	18
项目 10: 歌手大赛系统	20
10. 试题编号: 1-10-1 《歌手大赛系统》关键算法	20
项目 11: OJ 系统	22
11. 试题编号: 1-11-1 《OJ 系统》题库关键算法	22
项目 12: 字符处理系统	23
12. 试题编号: 1-12-1 《字符处理系统》关键算法	23
项目 13: 动物园管理系统	25
13. 试题编号: 1-13-1 《动物园管理系统》关键算法	25
项目 14: 中学生数学辅助学习系统	27
14. 试题编号: 1-14-1 《中学生数学辅助学习系统》关键算法	27
项目 15: 成绩分析系统	28
15. 试题编号: 1-15-1 《成绩分析系统》关键算法	28
模块二 数据库设计应用	30
项目 1: 银行信贷管理系统	30
1. 试题编号: 2-1-1 《银行信贷管理系统》客户管理模块	30
项目 1: 银行信贷管理系统	32
2. 试题编号: 2-1-2 《银行信贷管理系统》台帐管理模块	32
项目 2: 某电子商务网站	34
3. 试题编号: 2-2-1 《某电子商务网站》产品管理模块	34
项目 3: 学生食堂信息管理系统	35

4. 试题编号: 2-3-1 《学生食堂信息管理系统》饭卡管理模块.....	35
项目 4: 建设用地信息系统.....	37
5. 试题编号: 2-4-1 《建设用地信息系统》行政管理模块.....	37
项目 4: 建设用地信息系统.....	39
6. 试题编号: 2-4-2 《建设用地信息系统》地块管理模块.....	39
项目 5: 学生管理系统.....	40
7. 试题编号: 2-5-1 《学生管理系统》学生管理模块.....	40
项目 5: 学生管理系统.....	42
8. 试题编号: 2-5-2 《学生管理系统》成绩管理模块.....	42
项目 6: 音乐网站.....	44
9. 试题编号: 2-6-1 《音乐网站》项目歌手管理模块.....	44
项目 7: 网上商店.....	45
10. 试题编号: 2-7-1 《网上商店》商品管理模块.....	45
项目 7: 网上商店.....	47
11. 试题编号: 2-7-2 《网上商店》订单管理模块.....	47
项目 8: 图书管理信息系统.....	49
12. 试题编号: 2-8-1 《图书管理信息系统》借阅管理模块.....	49
项目 9: 起航自学考试网.....	51
13. 试题编号: 2-9-1 《起航自学考试网》课程管理模块.....	51
项目 10: 课程信息系统.....	52
14. 试题编号: 2-10-1 《课程信息系统》教师管理模块.....	52
项目 11: 仓库管理系统.....	54
15. 试题编号: 2-11-1 《仓库管理系统》货物管理模块.....	54
二、专业核心技能.....	56
模块三 大数据 HADOOP 技术应用.....	56
项目 1: 客户偏好电影推荐分析.....	56
1. 试题编号: 3-1-1 《客户偏好电影推荐分析》综合应用.....	56
项目 2: 网约车定位信息统计.....	57
2. 试题编号: 3-2-1 《网约车定位信息统计》综合应用.....	57
项目 3: 酒店客服出租率统计.....	59
3. 试题编号: 3-3-1 《酒店客服出租率统计》综合应用.....	59
项目 4: 薪水支出统计.....	60
4. 试题编号: 3-4-1 《薪水支出统计》统计汇总实现.....	60
项目 5: 股票日线数据清洗.....	62
5. 试题编号: 3-5-1 《股票日线数据清洗》数据清洗.....	62
项目 6: 学生消费数据清洗.....	63
6. 试题编号: 3-6-1 《学生消费数据清洗》去重操作实现.....	63
项目 7: 会员状况统计分析.....	65
7. 试题编号: 3-7-1 《会员状况统计分析》综合应用.....	65
项目 8: 金融风控分析.....	66
8. 试题编号: 3-8-1 《金融风控分析》综合应用.....	66
项目 9: 店铺零售业绩统计分析.....	68
9. 试题编号: 3-9-1 《店铺零售业绩统计分析》综合应用.....	68
项目 10: 面积统计分析.....	70

10. 试题编号: 3-10-1 《面积统计分析》综合应用.....	70
模块四 大数据技术应用	73
项目 1: 北京空气质量监测数据获取与分析.....	73
1. 试题编号: 4-1-1 北京空气质量监测数据获取与分析.....	73
项目 2: 58921 电影票房数据获取与分析.....	76
2. 试题编号: 4-2-1 58921 电影票房数据获取与分析.....	76
项目 3: 地表水水质监测数据获取与分析.....	79
3. 试题编号: 4-3-1 地表水水质监测数据获取与分析.....	80
项目 4: 空气吸收剂量率数据获取与分析.....	83
4. 试题编号: 4-4-1 空气吸收剂量率数据获取与分析.....	83
项目 5: 湖南城市 AQI 数据获取与应用.....	87
5. 试题编号: 2-5-1 湖南城市 AQI 数据获取与应用.....	87
项目 6: 地震历史数据获取与分析.....	90
6. 试题编号: 4-6-1 地震历史数据获取与分析.....	90
项目 7: 上海空气质量监测数据获取与分析.....	93
7. 试题编号: 4-7-1 上海空气质量监测数据获取与分析.....	94
项目 8: 杭州空气质量监测数据获取与分析.....	97
8. 试题编号: 4-8-1 杭州空气质量监测数据获取与分析.....	97
项目 9: 辐射监测数据获取与分析.....	100
9. 试题编号: 4-9-1 辐射监测数据获取与分析.....	100
项目 10: 广东城市 AQI 数据获取与应用.....	104
10. 试题编号: 4-10-1 广东城市 AQI 数据获取与应用.....	104

湖南生物机电职业技术学院软件技术专业

学生专业技能考核题库

一、专业基本技能

模块一 程序设计应用

项目 1: 小学生数学辅助学习系统

1. 试题编号: 1-1-1 《小学生数学辅助学习系统》关键算法

(1) 任务描述

随着社会的发展及人们对小学阶段的教育重视程度在不断提高, A 公司决定开发一套小学生数学辅助学习系统, 通过完成趣味试题, 采用游戏通关的方式, 帮助小学生掌握数学里的基本概念和计算方法。

任务一: 实现统计奇数个数的关键算法 (30 分)

通过键盘输入 2 个正整数, 统计并输出这 2 个数之间奇数的个数。例如, 输入 10 和 20, 奇数的个数为 5。

注意: 使用循环和分支结构语句实现。

任务二: 实现乘法口诀助记功能的关键算法 (30 分)

选择乘法口诀助记功能, 输出阶梯形式的 9*9 乘法口诀表, 如图 1.1.1 所示。

```
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
```

图 1.1.1 乘法口诀表

注意: 使用循环结构语句实现。

任务三: 实现求水仙花数的关键算法 (30 分)

要求输出所有的“水仙花数”。所谓“水仙花数”是指一个三位的整数, 其各位数字立方和等于该数本身。例如: 153 是一个“水仙花数”, 因为 $153=1^3+5^3+3^3$ 。

注意: 使用循环结构实现。

作品提交要求:

①请建立以“考生号_题号”命名的成果文件夹, 所有提交文件均放在该目录下。例如: 144115040001_J1_1;

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 task1、task2、task3, 将三个任务的源代

码、编译后的文件分别保存至对应的文件夹；

④将成果文件夹压缩打包，按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分（如一行编写两个语句）、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完

			为止。
--	--	--	-----

项目 2: 帮你算系统

2.试题编号: 1-2-1 《帮你算系统》关键算法

(1) 任务描述

随着网络的不断发展,我们每天接触的新鲜事物都在不断增加,处在这一个信息量大爆炸的时代,我们的时间就尤为重要,为了帮一些人解决时间不充裕的问题,处于创业的某公司准备开发一套“帮你算”系统,用来解决生活中那些简单、繁琐的数学问题。

任务一: 实现平均成绩和总成绩计算功能的关键算法 (30 分)

程序中提示用户录入该名学生的 5 门课程成绩,并统计 5 门课程的平均分和总分。

注意:使用数组存储学生各门课程成绩并输出该名学生的平均分和总分。

- ①定义数组 scores 存储 5 门课程成绩。
- ②定义 double 类型变量 totalScore 存储学生总成绩。
- ③定义 double 类型变量 avgScore 存储学生平均成绩。

任务二: 实现阶乘计算功能关键算法 (30 分)

输入一个整数 n, 计算并输出他的阶乘。

注意: 使用循环结构实现。

任务三: 实现从 1 到 n 的偶数和的关键算法 (30 分)

要求: 输入一个整数 n,利用循环计算该数列的偶数和。

注意: 使用分支结构实现

作品提交要求:

- ①请建立以“考生号_题号”命名的成果文件夹,所有提交文件均放在该目录下。例如: 144115040001_J1_1;
- ②分别将每个任务的代码按照题目所示要求书写到当前类中。
- ③在成果文件夹中创建三个文件夹 task1、task2、task3, 将三个任务的源代码、编译后的文件分别保存至对应的文件夹;
- ④将成果文件夹压缩打包,按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上, 内存 4G 以上, Win7/win10 操作系统	用于软件开发和软件部署, 每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量
180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分 (如一行编写两个语句)、使用空行不规范扣 1 分, 扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释, 但注释不规范扣 1 分, 扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范, 为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分, 扣完为止。

项目 3: 网络选拔赛题库系统

3. 试题编号: 1-3-1 《网络选拔赛题库系统》关键算法

(1) 任务描述

随着网络的普及, 许多比赛开始采用网络选拔赛的模式。某大赛组委会决定开发一个网络选拔赛题库系统, 实现该系统需要完成以下任务。

任务一: 实现统计元音字母关键算法 (30 分)

输入一个字符串统计每个元音字母 'a' 在字符串中出现的次数。

对于结果输出格式如下:

a:num (a 的个数)

例如输入:

aeioubbcccc

输出:

a:1

注意：使用分支语句实现。

任务二：实现大家来猜数关键算法（30 分）

自定义一个整数，请用户键盘输入猜想的数。比较两数大小，告诉用户猜大了还是小了，最多连续猜 5 次。如果用户猜对了，用户获胜；否则，输出设置的整数。

注意：使用循环语句和分支结构实现

任务三：实现 $2^x \bmod n = 1$ 关键算法（30 分）

给你一个数字 n ，找到满足 $2^x \bmod n = 1$ 的最小值 x ，如果 x 存在，则输出“ $x =$ 多少”，否则输出“ x 没有对应值”，您需要用真实的 x 和 n 的值来替代字符串中的变量。

例如输入 5，输出答案为 $x = 4$ 。

注意：使用分支结构实现

作品提交要求：

①请建立以“考生号_题号”命名的成果文件夹，所有提交文件均放在该目录下。例如：144115040001_J1_1；

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 task1、task2、task3，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；

④将成果文件夹压缩打包，按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 任务一评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
---	-----------	-----	-----------------------

2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分 (如一行编写两个语句)、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。

项目 4: 图形体积计算系统

4.试题编号: 1-4-1 《图形体积计算系统》关键算法

(1) 任务描述

图形在我们的生活中无处不在，看看我们的周围，到处都是图形的缩影、例如空调是长方形、水瓶瓶盖是圆形，这些图形的计算对于我们的土木工程师来说，那可是非常重要的啊，所以某公司开发出一套图形面积计算系统，帮助那些工程师们更好的计算。

任务一：实现计算体积关键算法 (30 分)

根据输入的半径值 r ，计算球的体积 v 。输出对应球的体积，计算结果保留三位小数。

$$\text{球体的体积 } v = (4.0 / 3) * \pi * r^3$$

注：PI 可以用 Math.PI 进行运算

例如：输入 2 输出 33.510

注意：使用公式完成。

任务二：实现坐标求长度关键算法 (30 分)

输入两点坐标 $(X1,Y1)$ ， $(X2,Y2)$ ，计算并输出两点间的距离。输入数据有多组，每组占一行，由 4 个实数组成，分别表示 $x1,y1,x2,y2$ ，数据之间用空格隔开。

例如输入：

1 3 4 6

则输出：

4.24

注意：结果保留两位小数。

任务三：实现图形面积大小比较关键算法（30 分）

按顺序输入正方形的边长（a），长方形的长（l）和宽（d），以及圆的半径（r），计算并比较它们哪个图形面积更大，输出面积最大的图形。

注：PI 可以用 Math.PI 进行运算

例如：输入 1 3 4 1，输出：长方形

作品提交要求：

①请建立以“考生号_题号”命名的成果文件夹，所有提交文件均放在该目录下。例如：144115040001_J1_1；

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 task1、task2、task3，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；

④将成果文件夹压缩打包，按照要求上传至服务器。

（2）实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

（3）考核时量

180 分钟。

（4）评分细则

表 2 任务一评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中

			出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则（10 分）

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分（如一行编写两个语句）、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。

项目 5：中国结图案打印系统

5.试题编号：1-5-1 《中国结图案打印系统》关键算法

(1) 任务描述

中国结是一种中国特有的手工编织工艺品，它身上所显示的情致与智慧正是中华古老文明的一个写照。它原本是由旧石器时代的缝衣打结，后推展至汉朝的仪礼记事，再演变成今日的装饰手艺。当代多用来装饰室内、亲友间的馈赠礼物及个人的随身饰物。因为其外观对称精致，可以代表中国悠久的历史，符合中国传统装饰的习俗和审美观念，故命名为中国结。现在 A 公司想要开发一个中国结图案打印系统，现在请你完成以下任务。

任务一：实现主结长度关键算法（30 分）

公司现在需要打印中国结的主结(位于中间，最大的那一个结)，为了打印出漂亮新颖的主结，于是设计打印主结的长度满足可以被 7 整除这个条件。现在公司需要统计某个范围内能被 7 整除的整数的个数，以及这些能被 7 整除的数的和。

从键盘上输入一个整数 N，输出 1~N 之间能被 7 整除的整数的个数，以及这些能被 7 整除的数的和。

注意：使用循环结构实现。

任务二：实现副结长度关键算法（30 分）

公司设计的中国结还需要副结(主结周围的结)，于是打算设计副结的长度满足是素数这个条件。现在公司需要统计出某个范围内哪些数是素数。

从键盘上输入一个整数 N，输出该数是否为素数。

注意：使用循环结构实现。

任务三：实现打印中国结图案关键算法（30 分）

中国结的形状各种各样，现公司需要设计一个打印三角形中国结的方法。

从键盘输入一个整数 N，打印出如图所示的三角形。

例如输入整数 5，则屏幕输出如下三角形。

*

现要求输入整数为 7，在屏幕中输出相应的三角形。
 注意：使用循环结构语句实现。

作品提交要求：

- ①请建立以“考生号_题号”命名的成果文件夹，所有提交文件均放在该目录下。例如：144115040001_J1_1；
- ②分别将每个任务的代码按照题目所示要求书写到当前类中。
- ③在成果文件夹中创建三个文件夹 task1、task2、task3，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；
- ④将成果文件夹压缩打包，按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中

			出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则（10 分）

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分（如一行编写两个语句）、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。

项目 6：智能统计系统

6.试题编号：1-6-1 《智能统计系统》关键算法

(1) 任务描述

生活中在处理各个问题的时候总是会离不开统计，例如统计学生的个数，统计火车买票人数，统计今天是该年的第几天等，所以某团队开发出一套统计系统，用来进行各类统计。

任务一：实现统计不同年份之间闰年个数关键算法（30 分）

从键盘上输入 2 个年份值，输出 2 个年份之间闰年的个数。

说明：闰年只需要满足以下条件之一即可：

1. 当前年份是 4 的倍数,而不是 100 的倍数.
2. 年份是 400 的倍数

注意：使用循环结构和分支结构语句实现。

任务二：实现统计纸片对折关键算法（30 分）

假设一张足够大的纸，纸张的厚度为 0.5 毫米。请问对折多少次以后，可以达到珠穆朗玛峰的高度(最新数据：8844.43 米)。请编写程序输出对折次数。

注意：使用循环结构语句实现，直接输出结果不计分。

任务三：实现统计同构数关键算法（30 分）

编写程序输出 2~99 之间的同构数。同构数是指这个数为该数平方的尾数，例如 5 的平方为 25，6 的平方为 36，25 的平方为 625，则 5、6、25 都为同构数。

注意：使用循环结构实现。

作品提交要求：

①请建立以“考生号_题号”命名的成果文件夹，所有提交文件均放在该目录下。例如：144115040001_J1_1；

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 task1、task2、task3，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；

④将成果文件夹压缩打包，按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分（如一行编写两个语句）、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。

项目 7：儿童智力游戏

7.试题编号：1-7-1 《儿童智力游戏》关键算法

(1) 任务描述

A 公司是专门的儿童网络游戏公司，现在公司正在开发几款智力游戏，其中需要设计几个算法模型。

任务一：实现积木游戏功能关键算法（30 分）

积木是小孩子最爱玩的游戏，但是因为小孩子的好奇心（比如误食积木等）导致家长们越不愿意让孩子去玩积木，为了解决这个问题 TX 公司开发了一套 VR 积木游戏，你要做的是将用户堆好的积木在屏幕中显示出来。

比如：输入积木的长 L 为 4 和宽 w 为 7,输出以下图形

```
*****
*****
*****
*****
```

注意：使用循环结构语句实现。

任务二：实现抓娃娃游戏功能关键算法（30 分）

请在娃娃机里放十个娃娃，每个娃娃对应一个数字，该数字表示娃娃的大小。要求通过计算能输出最大的娃娃对应的数字，你可以这样做：

- ① 定义一个大小为 5 的整形数组 arr；
- ② 从键盘输入 5 个整数，放置到数组 arr 中；
- ③ 输出数组 arr 中的最大值。

注意：使用数组、循环结构语句实现。

任务三：实现算数游戏功能关键算法（30 分）

游戏主要是这样的，分别统计正整数 1~n 之间偶数的个数和奇数的个数，并输出偶数多还是奇数多，现在要求你为 A 公司编写一个程序，计算结果,并显示输出。

注意:使用循环结构和分支结构实现.

作品提交要求：

- ①请建立以“考生号_题号”命名的成果文件夹，所有提交文件均放在该目录下。例如：144115040001_J1_1；
- ②分别将每个任务的代码按照题目所示要求书写到当前类中。
- ③在成果文件夹中创建三个文件夹 task1、task2、task3，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；
- ④将成果文件夹压缩打包，按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台

3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量
180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分 (如一行编写两个语句)、使用空行不规范扣 1 分, 扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释, 但注释不规范扣 1 分, 扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范, 为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分, 扣完为止。

项目 8: 商品销售系统

8.试题编号: 1-8-1 《商品销售系统》关键算法

(1) 任务描述

随着网络和信息化的发展, 电子商务越来越受到人们的欢迎。商品销售系统是电子商务中非常重要的业务支撑系统, 它能够为企业和商家提供充足的信息和快捷的查询手段, 能够让企业和商家了解自己的经营业绩、销售数据。现在需要完成以下任务来实现商品销售系统。

任务一: 实现打折功能关键算法 (30 分)

编写程序计算购买图书的总金额：用户输入图书的定价和购买图书的数量，并分别保存到一个 **double** 和一个 **int** 类型的变量中，然后根据用户输入的定价和购买图书的数量，计算购书的总金额并输出。其中，图书销售策略为：正常情况下按 9 折出售，购书数量超过 10 本打 8.5 折，超过 100 本打 8 折。

要求：使用分支结构实现上述程序功能。

任务二：实现查询功能关键算法（30 分）

已知 5 种商品的编号保存在一个 **int** 类型数组 **a** 中，根据用户输入的商品编号 **k**，查询该种商品是否存在，存在即显示其序号，否则显示“不存在”。数组 **a** 的元素值可以初始化或从键盘输入。

要求：使用循环结构语句实现，直接输出结果不计分。

任务三：实现商品总价和积分计算关键算法（30 分）

编写程序输入 5 个商品的价格,并录入到数组中,统计 5 个商品的总价,并输出对应积分,积分规则为：积分 = 总价 * 10

比如：录入 10,20,30,40,50. 输出 总消费：150 元,总积分：1500 积分。

要求：使用循环结构和数组实现。

作品提交要求：

①请建立以“考生号_题号”命名的成果文件夹，所有提交文件均放在该目录下。例如：144115040001_J1_1；

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 **task1**、**task2**、**task3**，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；

④将成果文件夹压缩打包，按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 任务一评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分 (如一行编写两个语句)、使用空行不规范扣 1 分, 扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释, 但注释不规范扣 1 分, 扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范, 为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分, 扣完为止。

项目 9: 幼师培训系统

9. 试题编号: 1-9-1 《幼师培训系统》关键算法

(1) 任务描述

幼儿教师是教师队伍中不可忽略的重要力量。她们主要以女性教育者为主, 负责教育学龄前儿童也就是幼儿。幼儿教师主要对幼儿进行启蒙教育帮助他们获得有益的学习经验, 促进其身心全面和谐发展。幼儿教师在教育过程中的角色决不仅仅是知识的传递者, 而且是幼儿学习活动的支持者、合作者、引导者。本系统用于培训幼师们与小朋友之间的游戏交互能力, 在游戏中可以帮助小朋友们的成长。为实现该系统, 需要完成以下任务。

任务一: 实现整除判断游戏功能关键算法 (30 分)

整除判断游戏能显著提高小朋友的逻辑思维能力, 问题要求如下:

- 能同时被 3、5 整除
- 只能被 3 整除
- 只能被 5 整除
- 不能被 3、5 任一个整除

输入一个整数, 输出满足对应条件的结果。

要求: 使用分支结构语句实现。

任务二: 实现小猴吃桃游戏功能关键算法 (30 分)

有一只猴子摘了一堆桃子, 每天吃掉一半桃子还不过瘾再多吃一个, 吃到第

十天的时候发现还剩下一个桃子，问猴子第一天一共摘了多少个桃子？

规律：

第 10 天：桃子数是：1

第 9 天：桃子数是： $(1+1)*2=4$

第 8 天：桃子数是： $(4+1)*2=10$

...

总结：前一天的桃子数 = (后一天的桃子数 + 1) * 2。

要求：使用循环结构语句实现。

任务三：实现数一数游戏关键算法（30 分）

输入一串字符，请问该组字符串中包含多少个数字'b'，如果没有则输出 0。

要求：使用循环结构语句实现。

作品提交要求：

①请建立以“考生号_题号”命名的成果文件夹，所有提交文件均放在该目录下。例如：144115040001_J1_1；

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 task1、task2、task3，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；

④将成果文件夹压缩打包，按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 任务一评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分 (如一行编写两个语句)、使用空行不规范扣 1 分, 扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释, 但注释不规范扣 1 分, 扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范, 为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分, 扣完为止。

项目 10: 歌手大赛系统

10.试题编号: 1-10-1 《歌手大赛系统》关键算法

(1) 任务描述

A 学校即将举行“校园歌手大赛”, 为了快速准确的分析选手的得分情况。A 学校决定开发一个歌手大赛系统。为了实现该系统的功能, 请完成以下三个任务。

任务一: 实现计算平均分功能关键算法 (30 分)

小明今天参加了“校园歌手大赛”, 现场有 4 名评委, 你能帮助小明快速的根据 4 名评委的评分算出平均分和总分吗?

输入说明: 输入 4 个整数分数。请按照题目要求计算平均分和总分

注意: 使用循环和数组实现。

任务二: 实现计算最小值功能关键算法 (30 分)

求 4 个整数中最小的数。每一个整数都独立看成一个数, 比如, 四个分数分别是 20, 91, 13, 71 那么, 最小的数就是 13。(每个数 均小于 100)

说明: 首先输入 4 个整数分数, 然后请输出最小的数。

注意: 使用循环实现。

任务三: 实现分数排序功能关键算法 (30 分)

小明的老师今天教了冒泡排序, 但是小明在网上找到一种更加方便的排序, 但是小明不会使用, 你能帮帮他吗? 现要求输入 4 个整数分数并录入数组, 输出从小到大的排序结果。

注意: 按照题目描述完成程序, 使用数组的 `sort` 方法实现。

作品提交要求:

①请建立以“考生号_题号”命名的成果文件夹, 所有提交文件均放在该目录下。例如: 144115040001_J1_1;

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 **task1**、**task2**、**task3**，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；

④将成果文件夹压缩打包，按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分（如一行编写两个语句）、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。

3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。
---	----------------	-----	---

项目 11: OJ 系统

11.试题编号: 1-11-1 《OJ 系统》题库关键算法

(1) 任务描述

在线评判系统(简称 OJ,Online Judge)指在线用来评判程序的正确性、时间与效率空间效率的评判系统。现需要为特定题目设计正确算法以便扩充题库,请完成以下任务。

任务一: 实现邮箱验证关键算法 (30 分)

提示用户输入一个邮箱字符串,如果用户输入的邮箱字符串的长度大于 3,并且邮箱中包含 '@' 字符,即认为该邮箱为合法邮箱,如果邮箱格式不合法,需对用户进行相应提示。

比如:

输入 “zhangsang@qq.com”, 该邮箱为合法邮箱.

输入 “lisi163.com”, 该邮箱不合法.

注意:使用分支结构实现

任务二: 实现运动会晋级关键算法 (30 分)

学校举行 5000m 长跑运动大会,运动大会规则分为男女俩组,男子组晋级规则为时间小于 17 分,女子组晋级规则为时间小于 20 分.现需要 OJ 系统提供程序来判断运动员是否晋级.

程序要求,比如

请输入运动员性别: 男

请输入运动员成绩: 16

输出: 恭喜你,成功晋级

注意:使用分支结构实现

任务三: 实现数字校验规则关键算法 (30 分)

输出 100 内能被 3 整除但是不能被 4 整除的所有奇数.

注意:使用分支结构和循环结构实现

作品提交要求:

①请建立以“考生号_题号”命名的成果文件夹,所有提交文件均放在该目录下。例如: 144115040001_J1_1;

②分别将每个任务的代码按照题目所示要求书写到当前类中.

③在成果文件夹中创建三个文件夹 task1、task2、task3,将三个任务的源代码、编译后的文件分别保存至对应的文件夹;

④将成果文件夹压缩打包,按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间

			距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量
180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分（如一行编写两个语句）、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。

项目 12: 字符处理系统

12.1 试题编号: 1-12-1 《字符处理系统》关键算法

(1) 任务描述

在印刷生产中，要求对打印或印刷的数字字符进行实时识别校验，如卡号和密码，然后把打印或印刷错误的字符串剔除。目前检查字符串的主要方法是通过

人工目视检查，存在效率低、准确度不高的问题。因此开发一套字符处理系统就显得尤为重要，目前实现字符处理系统还需要完成如下任务。

任务一：实现用户登录关键算法（30 分）

定义一个字符串来变量存放用户输入的密码，如果用户输入的密码等于系统密码，则输出欢迎登录的提示信息；否则提示重新输入密码，输入 3 次错误则结束程序。

注意：使用循环结构和分支结构实现。

任务二：实现找最小值关键算法（30 分）

循环录入 5 个不同的整数，找出其中最小的数，并输出该值

要求：使用循环结构和分支结构实现。

任务三：实现统计字符数量功能关键算法（30 分）

对于给定的一个字符串，统计其中字母'z'出现的次数。

要求：使用循环结构和分支结构实现。

作品提交要求：

①请建立以“考生号_题号”命名的成果文件夹，所有提交文件均放在该目录下。例如：144115040001_J1_1；

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 task1、task2、task3，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；

④将成果文件夹压缩打包，按照要求上传至服务器。

（2）实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

（3）考核时量

180 分钟。

（4）评分细则

表 2 任务一评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则（30 分）

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分（如一行编写两个语句）、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。

项目 13: 动物园管理系统

13.试题编号: 1-13-1 《动物园管理系统》关键算法

(1) 任务描述

动物园内饲养了大量不同种类的动物，因此对这些动物的生活场地的建设及食物投放的管理工作量非常大。现在 A 动物园需要设计并实现一套动物园管理系统，提高管理效率，请完成以下任务：

任务一：实现饲养功能关键算法 (30 分)

动物园现准备统计饲养动物的食量，动物园为了方便统计，从编号 1 开始的笼子就喂食肉类 1 斤，编号 2 的笼子喂食 2 斤，现已知动物园有 100 个笼子。请设计一个程序帮助饲养员计算一天喂食的数量。

要求：用循环语句实现。

任务二：实现找找哪个动物最大问题关键算法 (30 分)

动物园里分别新来了三只动物，乌龟，大象，熊猫，那么你能使用程序计算哪只动物年龄最大么？

从键盘输入三个整数，比如 400,10,2。输出乌龟年龄最大

要求：用分支语句实现。

任务三：实现人工湖关键算法 (30 分)

现在，动物园想在新建一个三角形的人工湖，一是为了养鱼美观，二是可以循环水资源。

从键盘输入三条边 A、B、C 的边长，请编程判断能否组成一个三角形。

要求: A,B,C <1000, 如果三条边长 A、B、C 能组成三角形的话, 输出 YES, 否则 NO,组成三角形的条件为,任意俩边之和大于第三边

作品提交要求:

①请建立以“考生号_题号”命名的成果文件夹, 所有提交文件均放在该目录下。例如: 144115040001_J1_1;

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 task1、task2、task3, 将三个任务的源代码、编译后的文件分别保存至对应的文件夹;

④将成果文件夹压缩打包, 按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上, 内存 4G 以上, Win7/win10 操作系统	用于软件开发和软件部署, 每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中出现了无用的循环、分支、循序结构扣 1 分, 扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能, 否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分（如一行编写两个语句）、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。

项目 14：中学生数学辅助学习系统

14.试题编号：1-14-1 《中学生数学辅助学习系统》关键算法

(1) 任务描述

由于中学数学是培养数学思维的基础阶段，为了让学生打造一个坚实的数学基础，A 学校决定开发一个中学生数学辅助学习系统，通过完成趣味试题，采用游戏通关的方式，帮助中学生初步掌握解简单应用题的方法和步骤，并会列出解简单的应用题。

任务一：实现趣味数字统计的关键算法（30 分）

现准备等差数列 1,2,3,4...100,等差数列求和公式为(首项 + 尾项) * 个数 / 2, 但为了推导公式实现过程,现在需要学生计算从 1 累加到 100 的和来验证公式是否正确,最后计算该等差数列的平均值.

注意：用循环语句实现。

任务二：实现找找谁最大的关键算法（30 分）

现需要使用程序循环录入 5 个同学的年龄,通过两数比较的方式计算出年龄最大的同学的年龄.

注意：用循环语句实现。

任务三：实现统计特定数字的关键算法（30 分）

现需要使用程序找出 1000 - 9999 之间的所有回文数,回文数是指正序（从左向右）和倒序（从右向左）读都是一样的整数。

比如 1221,3663

注意：用循环语句和分支结构实现。

作品提交要求：

①请建立以“考生号_题号”命名的成果文件夹，所有提交文件均放在该目录下。例如：144115040001_J1_1;

②分别将每个任务的代码按照题目所示要求书写到当前类中。

③在成果文件夹中创建三个文件夹 task1、task2、task3，将三个任务的源代码、编译后的文件分别保存至对应的文件夹；

④将成果文件夹压缩打包，按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
----	------------	------------	----

1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上，内存 4G 以上，Win7/win10 操作系统	用于软件开发和软件部署，每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量
180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分（如一行编写两个语句）、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。

项目 15: 成绩分析系统

15.试题编号: 1-15-1 《成绩分析系统》关键算法

(1) 任务描述

对学生成绩进行统计和数据分析可以发现学生对知识的掌握情况,以便教师根据分析的结果调整教学内容和重难点,现在需要完成以下任务来实现成绩分析系统。

任务一：实现成绩等级划分功能关键算法（30 分）

输入一个百分制的成绩 t , 将其转换成对应的等级然后输出, 具体转换规则如下:

80~100 输出 A

60~79 输出 B

0~59 输出 C

其他 输出 “error”

注意:使用分支结构实现。

任务二：实现成绩统计求和功能关键算法（30 分）

现输入学生的 5 门功课成绩,保存到数组中,最后统计 5 门功课的总分并输出. 注意:使用循环结构实现.

任务三：求前 n 项之和功能关键算法（30 分）

多项式的描述如下: $1 + 1/2 + 1/3 + 1/4 + 1/5 + 1/6 + \dots$, 现在要求出该多项式的前 n 项的和。

要求: 使用循环结构。

作品提交要求:

- ①请建立以“考生号_题号”命名的成果文件夹, 所有提交文件均放在该目录下。例如: 144115040001_J1_1;
- ②分别将每个任务的代码按照题目所示要求书写到当前类中。
- ③在成果文件夹中创建三个文件夹 `task1`、`task2`、`task3`, 将三个任务的源代码、编译后的文件分别保存至对应的文件夹;
- ④将成果文件夹压缩打包, 按照要求上传至服务器。

(2) 实施条件

表 1 考点提供的主要设备及软件表

序号	场地、设备、软件名称	规格/技术参数、用途	备注
1	软件技术实训机房	测试场地	保证参考人员有足够间距
2	计算机	CPU 酷睿 i5 以上, 内存 4G 以上, Win7/win10 操作系统	用于软件开发和软件部署, 每人一台
3	Eclipse 4.7, idea 14 或以上	软件开发	参考人员自选一种开发工具
4	JDK 帮助文档中文版	帮助文档	参考人员可以使用帮助文档

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 任务一评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件, 扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分; 程序中

			出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 3 任务二评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 4 任务三评分细则 (30 分)

1	开发环境使用正确性	5 分	未按要求提交正确格式的源文件，扣 5 分。
2	程序设计合理性	15 分	程序中出现了没有使用的变量扣 1 分；程序中出现了无用的循环、分支、循序结构扣 1 分，扣完为止。
3	功能实现	10 分	按照任务要求实现相应功能，否则记 0 分。

表 5 职业素质评分细则 (10 分)

序号	评分项	分值	评分细则
1	代码书写格式规范	3 分	代码缩进不规范扣 1 分、方法划分不规范扣 1 分、语句结构不规范扣 1 分 (如一行编写两个语句)、使用空行不规范扣 1 分，扣完为止。
2	注释规范	2 分	整个项目没有注释扣 2 分、有注释，但注释不规范扣 1 分，扣完为止。
3	类名、变量名、方法名命名规范	5 分	命名规范，为满分。类名、变量名或方法名命名不规范或没有实际意义的每个扣 1 分，扣完为止。

模块二 数据库设计应用

项目 1: 银行信贷管理系统

1. 试题编号: 2-1-1 《银行信贷管理系统》客户管理模块

(1) 任务描述

《客户管理》模块的数据表字段名定义见表 2.1.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.1.1 字段名定义表

字段名	类型	备注	约束
cust_id	char(10)	客户编号	主键
cust_name	varchar(60)	客户名称	非空
legal_name	varchar(10)	法人代表	非空
reg_address	varchar(60)	注册地址	默认‘湖南’
post_code	char(6)	邮政编码	无
cust_tel	varchar(20)	联系电话	无

任务一：创建数据库 (10 分)

创建数据库 BankCreditLoanDB。

任务二：创建数据表（25 分）

根据表 2.1.1，创建数据表 T_customer_info。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

注册地址列(reg_address)默认值为‘湖南’。

任务四：数据操作（40 分）

用 SQL 语句完成如下操作：

①在 T_customer_info 表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）；

②查询所有客户的信息；

③删除表中联系电话为 0732 开头的记录；

④将客户名称为长沙红凯发展有限公司的法人代表改为李红；

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql”，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 1：银行信贷管理系统

2. 试题编号：2-1-2 《银行信贷管理系统》台帐管理模块

(1) 任务描述

《台帐管理》模块的数据表字段名定义见表 2.2.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.2.1 字段名定义表

字段名	类型	备注	约束
borrow_id	char(10)	借据号	主键
cust_id	char(10)	客户编号	非空
Agency_id	char(3)	客户经理编号	非空
pawn_goods_name	varchar(60)	押物名称	无
pawn_goods_num	int	押物数量	无
contract_date	datetime	签订日期	默认当前日期

任务一：创建数据库（10 分）

创建数据库 BankCreditLoanDB。

任务二：创建数据表（25 分）

根据表 2.2.1，创建数据表 T_impawn_info。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

签订日期列(contract_date)默认值为当前日期；

任务四：数据操作（40 分）

用 SQL 语句完成如下操作：

①在 T_impawn_info 表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）；

②查询所有借据的信息；

- ③删除客户编号为‘c001’的记录;
- ④将 T_impawn_info 表中的所有的押物数量增加 1;

任务五：代码规范（10分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db\”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。
数据库备份	5 分	没有备份数据库每项扣 5 分。
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。
代码数据库命名规范	2 分	数据库命名不规范扣 2 分。

数据表命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
字段命名规范	5 分	字段命名不规范每项扣 0.5 分，扣完为止。

项目 2：某电子商务网站

3.试题编号：2-2-1 《某电子商务网站》产品管理模块

(1) 任务描述

《产品管理》模块的数据表字段名定义见表 2.3.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.3.1 字段名定义表

字段名	类型	备注	约束
product_id	int	产品编号	主键 自增
category_id	int	产品类别 ID	唯一键
product_name	varchar(50)	产品名称	非空
price	money	产品价格	非空
remark	varchar(2000)	产品描述	无
register_date	datetime	录入时间	无

任务一：创建数据库（10 分）

创建数据库 ProductDB。

任务二：创建数据表（25 分）

创建数据表 T_product，其中表的产品 ID(product_id)列设置为标识列，自动从 1 开始增长。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

产品类别 ID 列(category_id)设置为唯一键；

任务四：数据操作（40 分）

用 SQL 语句查询出如下数据：

①在 T_product 表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）；

②查询产品编号为 2 的所有信息；

③删除产品类别 ID 为‘1001’的记录；

④将 T_product 表中的商品编号为 5 的商品价格上调 10%；

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a. SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql”，存放 SQL 脚本代码。

b. 数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 3：学生食堂信息管理系统

4. 试题编号：2-3-1 《学生食堂信息管理系统》饭卡管理模块

(1) 任务描述

《饭卡管理》模块的数据表字段名定义见表 2.4.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.4.1 字段名定义表

字段名	类型	备注	约束
-----	----	----	----

card_id(标识列)	int	饭卡编号	主键, 自增
student_id	char(10)	学生学号	唯一键
student_name	varchar(20)	学生姓名	非空
curr_money	money	饭卡余额	无
register_date	datetime	饭卡生成时间	默认为当前录入时间

任务一：创建数据库（10 分）

创建数据库 CardDB。

任务二：创建数据表（25 分）

创建数据表 T_card，其中表的饭卡编号(card_id)列设置为标识列，自动从 1 开始增长。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

学生学号列(student_id)设置为唯一键；

饭卡生成时间列(register_date)默认值为当前录入时间；

任务四：数据操作（40 分）

①在 T_card 表中插入 5 条测试数据(样本数据包含下面题目中使用的数据)；

②查询所有学生的饭卡信息；

③删除学生学号为‘s001’的记录；

④将 T_card 表中所有的饭卡余额增加 60；

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db\”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

实施条件要求见本模块附录 2。

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库, 扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分, 字段创建不符合要求每一项扣 2-3 分, 扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分, 关系创建不符合要求每一项扣 5 分, 扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分, 扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确, 无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分, 扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分, 扣完为止。

项目 4: 建设用地信息系统

5. 试题编号: 2-4-1 《建设用地信息系统》行政管理模块

(1) 任务描述

《行政管理》模块的数据表字段名定义见表 2.5.1。请按以下设计完成数据库创建、数据表创建和数据操作任务:

表 2.5.1 字段名定义表

字段名	类型	备注	约束
canton_id	int	行政区 ID	主键, 自增
canton_no	varchar(5)	行政区编号	唯一键
canton_name	varchar(50)	行政区名称	非空
canton_tot	float	行政区面积	无
branch	varchar(50)	直属部门	无
remark	text	备注	无

任务一: 创建数据库 (10 分)

创建数据库 Proj。

任务二: 创建数据表 (25 分)

创建数据表 T_canton_info, 其中表的行政区 ID (canton_id) 列设置为标识列, 自动从 1 开始增长。

任务三: 创建数据表间的关系及约束 (15 分)

为表设置主键, 主键命名为“pk_<表名>_<主键标识>”;

行政区编号列(canton_no)设置为唯一键;

任务四: 数据操作 (40 分)

用 SQL 语句执行以下操作:

①在 T_canton_info 表中插入 5 条测试数据 (样本数据包含下面题目中使用的数据);

②查询所有行政区 ID 的信息;

③删除行政区编号为‘c001’的记录;

④将 T_canton_info 表中行政区 ID 为 2 的直属部门改为教育部;

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql”，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db\”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。

范		
字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 4：建设用地信息系统

6.试题编号：2-4-2 《建设用地信息系统》地块管理模块

(1) 任务描述

《地块管理》模块的数据表字段名定义见表 2.6.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.6.1 字段名定义表

字段名	类型	备注	约束
plot_id	varchar(5)	地块编号	主键
proj_id	int	用地项目 ID	唯一键
plot_name	varchar(20)	地块名称	非空
total_area	float	地块面积	非空
purpose	varchar(50)	地块用途	默认商用
shape_type	varchar(50)	图形类型	无

任务一：创建数据库（10 分）

创建数据库 ContractionDB。

任务二：创建数据表（25 分）

创建数据表 T_plot_info。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

用地项目 ID 列(proj_id)设置为唯一键；

地块用途列(purpose)设置默认为商用；

任务四：数据操作（40 分）

用 SQL 语句执行以下操作：

①在 T_plot_info 表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）；

②查询出所有的地块基本信息并按地块面积升序排序；

③删除用地项目 ID 为 1001 的记录；

④将 T_plot_info 表中地块编号为‘p002’的地块用途改为工业用地；

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql，存放 SQL

脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db\”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 5：学生管理系统

7.试题编号：2-5-1 《学生管理系统》学生管理模块

(1) 任务描述

《学生管理》模块的数据表字段名定义见表 2.7.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.7.1 字段名定义表

字段名	类型	备注	约束
student_no	char(6)	学号	主键

student_name	char(20)	姓名	非空
sex	char(2)	性别	只能为‘男’或‘女’
birth	datetime	出生年月	非空
enter_date	int	入学年份	非空
address	varchar(50)	家庭住址	无

任务一：创建数据库（10 分）

创建数据库 Student。

任务二：创建数据表（25 分）

创建数据表 T_student_info

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

为性别列(sex)设置 check 约束只能为‘男’或‘女’；

任务四：数据操作（40 分）

用 SQL 语句完成如下操作：

①在 T_student_info 表中插入 5 条测试数据(样本数据包含下面题目中使用的数据)；

②查询出 1991 年出生的学生信息；

③删除姓李的学生信息；

④将 T_student_info 表中张三的出生年月改为‘1992-3-1’；

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db\”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	

3	SQL Server 2005 或以上、 Oracle 10g 或以上、 MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库 管理系统
---	--	---------	---------------------

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 5：学生管理系统

8.试题编号：2-5-2 《学生管理系统》成绩管理模块

(1) 任务描述

《成绩管理》模块的数据表字段名定义见表 2.8.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.8.1 字段名定义表

字段名	类型	备注	约束
Course_no	char(8)	课程编号	联合主键
Student_no	char(6)	学号	联合主键
Ordinary_scores	decimal(7,2)	平时成绩	非空
End_scores	decimal(7,2)	期末成绩	非空
Total_scores	decimal(7,2)	总评成绩	只能是 0-100 之间

任务一：创建数据库（10 分）

创建数据库 Student。

任务二：创建数据表（25 分）

创建数据表 T_student_scores。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

为总评成绩列(Total_scores)设置 check 约束分数只能是 0-100 之间；

任务四：数据操作（40 分）

用 SQL 语句完成如下操作：

①在 T_student_scores 表中插入 5 条测试数据(样本数据包含下面题目中使

用的数据);

- ②查询所有学生成绩信息并按总评成绩的升序排序;
- ③删除学号's001'的课程编号为'c001'的成绩信息;
- ④将 T_student_scores 表中'c002'课程的总评成绩减 2 分;

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql”，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db\”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。
数据库备份	5 分	没有备份数据库每项扣 5 分。
数据库管理系统配	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5

置与使用		分。	
代 码 规 范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 6: 音乐网站

9. 试题编号: 2-6-1 《音乐网站》项目歌手管理模块

(1) 任务描述 《歌手管理》模块的数据表字段名定义见表 2.9.1。请按以下设计完成数据库创建、数据表创建和数据操作任务:

表 2.9.1 字段名定义表

字段名	类型	备注	约束
Id	int	歌曲编号	主键, 自增
Type	varchar (50)	歌曲类型	非空
Name	varchar (50)	歌手名	非空
Sex	varchar (2)	歌手性别	只能为男或女, 默认为男
Profile	varchar (50)	歌手简介	无
Head	varchar (50)	歌手头像	非空

任务一: 创建数据库 (10 分)

创建数据库 SingerDB。

任务二: 创建数据表 (25 分)

根据表 2.9.1, 创建数据表 tbl_Singer。

任务三: 创建数据表的约束 (15 分)

为表设置主键, 主键命名为“pk_<表名>_<主键标识>”;

为歌手性别列(Sex)设置 check 约束只能为男或女;

歌手性别默认为男。

任务四: 数据操作 (40 分)

用 SQL 语句完成如下操作:

①在表中插入 5 条测试数据 (样本数据包含下面题目中使用的数据)。

②查询所有的歌手信息;

③修改歌手名为李明的性别为女;

④删除歌手简介为空的记录。

任务五: 代码规范 (10 分)

数据库命名规范, 数据表命名规范, 字段命名规范。

作品提交要求:

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件:

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹, 存放所有答题文件, 例如:

“340103*****_J2_1\”

②创建答题文件

a. SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql”，存放 SQL 脚本代码。

b. 数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 7：网上商店

10.试题编号：2-7-1 《网上商店》商品管理模块

(1) 任务描述

《商品管理》模块的数据表字段名定义见表 2.10.1。请按以下设计完成数据

库创建、数据表创建和数据操作任务：

表 2.10.1 字段名定义表

字段名	类型	备注	约束
goods_id	int	商品编号	主键, 自增
supp_code	varchar(20)	供应商编号	非空
goods_name	varchar(50)	商品名称	非空
goods_price	decimal(10,2)	商品价格	非空
goods_adress	varchar(50)	生产地	无
register_date	datetime	入库时间	默认为当前系统时间

任务一：创建数据库（10 分）

创建数据库 Stores。

任务二：创建数据表（25 分）

创建数据表 T_goods。

将 T_goods 表的商品编号 (Goods_id)列设置为标识列, 自动从 1 开始增长。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键, 主键命名为“pk_<表名>_<主键标识>”;

为入库时间列(register_date)设置默认为当前系统时间。

任务四：数据操作（40 分）

用 SQL 语句完成如下操作：

①在表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）。

②查询所有商品的平均价格;;

③将商品名为“惠普笔记本”的价格下调 10%;

④删除生产地为台湾的记录。

任务五：代码规范（10 分）

数据库命名规范, 数据表命名规范, 字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹, 存放所有答题文件, 例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件, 如: “340103*****_J2_1\project.sql, 存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹, 如: “340103*****_J2_1\db\”, 存放数据库备份文件, 它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包, 形成“考生号_题号.rar”文件, 如：

“340103*****_J2_1.rar”, 将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上, 内存 4G 以上, win7/win10 操作系统	用于软件开发和软件部署, 每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库, 扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分, 字段创建不符合要求每一项扣 2-3 分, 扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分, 关系创建不符合要求每一项扣 5 分, 扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分, 扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确, 无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分, 扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分, 扣完为止。

项目 7: 网上商店

11. 试题编号: 2-7-2 《网上商店》订单管理模块

(1) 任务描述

《订单管理》模块的数据表字段名定义见表 2.11.1。请按以下设计完成数据库创建、数据表创建和数据操作任务:

表 2.11.1 字段名定义表

字段名	类型	备注	约束
order_id	int	订单号码	主键, 自增
goods_id	int	商品编号	唯一键
goods_name	varchar(50)	商品名称	非空
quantity	int	订购数量	非空
Total_price	decimal(10,2)	总价	非空
order_date	datetime	订货日期	非空

任务一: 创建数据库 (10 分)

创建数据库 Stores。

任务二: 创建数据表 (25 分)

创建数据表 T_orders ，将 T_orders 表的商品编号 (order_id)列设置为标识列，自动从 1 开始增长。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

为商品编号列(goods_id)设置为唯一键。

任务四：数据操作（40 分）

用 SQL 语句完成如下操作：

①在表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）。

②查询所有商品的信息，并按总价的降序排序；

③将商品名为“惠普笔记本”的订购数量增加 10；

④删除 2019 年 3 月的订单。

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db\”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

（2）实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

（3）考核时量

180 分钟。

（4）评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代 码 规 范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 8：图书管理信息系统

12.试题编号：2-8-1 《图书管理信息系统》借阅管理模块

(1) 任务描述

《借阅管理》模块的数据表字段名定义见表 2.12.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.12.1 字段名定义表

字段名	类型	备注	约束
book_no	varchar(20)	书号	主键
book_name	varchar(50)	书名	非空
author	varchar(50)	作者	非空
price	decimal(10,2)	单价	非空
qty	int	库存册数	只能为 1 到 30
loan_qty	int	借阅次数	非空

任务一：创建数据库（10 分）

创建数据库 BookDB。

任务二：创建数据表（25 分）

创建数据表 T_book。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

为库存册数列(qty)设置只能为 1 到 30 的数。

任务四：数据操作（40 分）

用 SQL 语句完成如下操作：

①在表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）。

②查询出库存册数小于 5 册的图书信息；

③将书号为“b002”的单价下调 10%；

④删除借阅次数小于 10 的记录。

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 9：起航自学考试网

13.试题编号：2-9-1《起航自学考试网》课程管理模块

(1) 任务描述

《课程管理》模块的数据表字段名定义见表 2.13.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.13.1 字段名定义表

字段名	类型	备注	约束
course_id	varchar(20)	课程 ID	主键
course_name	varchar(20)	课程名称	非空
speciality_id	varchar(20)	专业 ID	非空
time	int	课程学时	非空
mark	int	课程学分	只能是 1 到 6 分之间

任务一：创建数据库（10 分）

创建数据库 SelfStudy。

任务二：创建数据表（25 分）

创建数据表 T_course。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

为课程学分列(mark)设置只能为 1 到 6 的数。

任务四：数据操作（40 分）

在表中插入以下数据，用作测试。

用 SQL 语句完成如下操作：

- ①在表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）。
- ②查询“电子政务”专业开设的课程；
- ③将课程名称为 SQLServer 数据库的课程学分改为 6 分；
- ④删除课程学分为 1 的记录。

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql”，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db\”，存放数据库备份文

件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

(2) 实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

(3) 考核时量

180 分钟。

(4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代码规范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

项目 10：课程信息系统

14.试题编号：2-10-1 《课程信息系统》教师管理模块

(1) 任务描述

《教师管理》模块的数据表字段名定义见表 2.14.1。请按以下设计完成数据库创建、数据表创建和数据操作任务：

表 2.14.1 字段名定义表

列名	类型	备注	约束
id	int	教师编号	主键

name	varchar (50)	姓名	非空
sex	varchar (50)	性别	默认男
course	varchar (50)	课程号	非空
department	varchar (50)	院系	无
title	varchar (50)	职称	只能为助教、讲师、副教授、教授

任务一：创建数据库（10 分）

创建数据库 TeacherDB。

任务二：创建数据表（25 分）

创建数据表 tbl_teacher_info。

任务三：创建数据表间的关系及约束（15 分）

为表设置主键，主键命名为“pk_<表名>_<主键标识>”；

为性别列(sex)设置默认值男。

为职称列(title)设置只能为助教、讲师、副教授、教授。

任务四：数据操作（40 分）

利用数据管理工具在表中插入以下数据，用作测试。

用 SQL 语句完成如下操作：

- ①在表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）；
- ②查询教师名为李磊的所有信息；
- ③修改教师名为李磊的课程号为 74822；
- ④删除教师编号为 1001 的记录。

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql”，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db\”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

（2）实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以	用于软件开发和软件部
---	-----	----------------------	------------

		上, win7/win10 操作系统	署, 每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、 Oracle 10g 或以上、 MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库 管理系统

- (3) 考核时量
180 分钟。
- (4) 评分细则

表 2 考核评价细则表

评价项	分值	评分细则	
数据库创建	10 分	没有成功创建数据库, 扣 5-8 分。	
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分, 字段创建不符合要求每一项扣 2-3 分, 扣完为止。	
约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分, 关系创建不符合要求每一项扣 5 分, 扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分, 扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确, 无法连接数据库扣 5 分。	
代 码 规 范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分, 扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分, 扣完为止。

项目 11: 仓库管理系统

15. 试题编号: 2-11-1 《仓库管理系统》货物管理模块

(1) 任务描述

《货物管理》模块的数据表字段名定义见表 2.15.1。请按以下设计完成数据库创建、数据表创建和数据操作任务:

表 2.15.1 字段名定义表

列名	类型	备注	约束
good_id	int	货物编号	主键
depot_id	int	仓库编号	唯一键
good_name	varchar(50)	货物名称	非空
good_type	varchar(50)	货物类型	非空
good_total	float	货物总量	非空
good_date	datetime	入库时间	默认为当前录入时间

任务一: 创建数据库 (10 分)

创建数据库 StoresControlDB。

任务二: 创建数据表 (25 分)

创建数据表 T_Good_info。

任务三: 创建数据表间的关系及约束 (15 分) 及约束

为表设置主键, 主键命名为“pk_<表名>_<主键标识>”;

为入库时间列(depot_id)设置唯一键;
为性别列(good_date)设置默认值为当前录入时间;

任务四：数据操作（40 分）

用 SQL 语句完成如下操作：

- ①在表中插入 5 条测试数据（样本数据包含下面题目中使用的数据）；
- ②查询仓库编号为 1001 的存放货物信息；
- ③将货物编号为 12 的入库时间改为 2020 年 8 月 10 日；
- ④删除货物类型为汽车用品的所有信息。

任务五：代码规范（10 分）

数据库命名规范，数据表命名规范，字段命名规范。

作品提交要求：

答案以“答题文件”的形式提交。请按以下要求创建答题文件夹和答题文件：

①创建答题文件夹

1 创建答题文件夹

创建以“考生号_题号”命名的文件夹，存放所有答题文件，例如：

“340103*****_J2_1\”

②创建答题文件

a.SQL 脚本文件

创建 project.sql 文件，如：“340103*****_J2_1\project.sql，存放 SQL 脚本代码。

b.数据库文件

创建 db 子文件夹，如：“340103*****_J2_1\db”，存放数据库备份文件，它

用于教师阅卷时还原数据库。

③提交答题文件

将“考生号_题号”文件夹打包，形成“考生号_题号.rar”文件，如：

“340103*****_J2_1.rar”，将该文件按要求进行上传。

（2）实施条件

表 1 考点提供的主要设备及软件表

1	计算机	CPU 酷睿 i5 以上，内存 4G 以上，win7/win10 操作系统	用于软件开发和软件部署，每人一台
2	Office	编写文档	
3	SQL Server 2005 或以上、Oracle 10g 或以上、MySQL 5.1 或以上	数据库管理系统	参考人员任选一种数据库管理系统

（3）考核时量

180 分钟。

（4）评分细则

表 2 考核评价细则表

评价项	分值	评分细则
数据库创建	10 分	没有成功创建数据库，扣 5-8 分。
数据表创建	25 分	数据表创建不成功每一项扣 3-5 分，字段创建不符合要求每一项扣 2-3 分，扣完为止。

约束及关系创建	15 分	约束创建不成功每一项扣 3-5 分，关系创建不符合要求每一项扣 5 分，扣完为止。	
数据访问	30 分	没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。	
数据库备份	5 分	没有备份数据库每项扣 5 分。	
数据库管理系统配置与使用	5 分	数据库服务器与管理工具配置不正确，无法连接数据库扣 5 分。	
代 码 规 范	数据库名命名规范	2 分	数据库名命名不规范扣 2 分。
	数据表名命名规范	3 分	数据表命名不规范每张表扣 1 分，扣完为止。
	字段名命名规范	5 分	字段名命名不规范每项扣 0.5 分，扣完为止。

二、专业核心技能

模块三 大数据 Hadoop 技术应用

项目 1: 客户偏好电影推荐分析

1.试题编号: 3-1-1 《客户偏好电影推荐分析》综合应用

(1) 任务描述

在电影网站发展初始阶段，电影推荐功能就应运而生。电影推荐根据热门数据及用户的个性化数据，为用户提供定制化的推荐。

现有用户电影评分数据，存储在/root/experiment/datas/user_movie.txt 文件中，字段包括：用户 ID (user_id)，电影 ID (movie_id)，用户评分 (user_score)；

有电影信息数据，存储在/root/experiment/datas/movie.txt 文件中，字段包括：电影 ID (movie_id)，电影名称 (movie_name)。

请使用 HiveQL 语句，完成以下功能。

- 1.以电影 ID 作为关联字段，关联电影评分数据与电影信息数据
- 2.根据关联数据表统计各电影的平均评分。

任务一：启动 Hive 操作窗口。

任务二：创建电影评分数据表 t_user_movie 并导入数据。

任务三：创建电影信息表 t_movie 并导入数据。

任务四：创建关联数据表 t_movie_join (需求字段 movie_id, movid_name, user_score)，以电影 ID 作为关联字段关联 t_user_movie 与 t_movie 并将关联数据插入 t_movie_join 表。

任务五：查询各电影的平均评分。

(2) 实施条件

抽测所需的软硬件如表 H1-4-1:

表 H1-4-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上，内存 16GB，硬盘空间 500G 以上)；该计算机上搭载了或者能够访问 Hadoop 平台，该 Hadoop 平台上已安装好 Hive 数仓。	每人一台

项目	基本实施条件	备注
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	远程访问工具, Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制, 评价内容包括工作任务、职业素养完成情况两个方面。其中, 工作任务完成质量占该项目总分的 90%, 职业素养占该项目总分的 10%。具体评价标准见下表:

表 H1-4-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注
工作任务 (90分)	启动 Hive 操作窗口	10 分	能够正确启动 hadoop 服务器, 能启动 Hive	1、考试舞弊、抄袭、没有按要求填写相关信息, 本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。
	创建电影评分数据表 t_user_movie 并导入数据	15 分	能够正确创建数据表, 并导入数据	
	创建电影信息表 t_movie 并导入数据	15 分	能够正确创建数据表, 并导入数据	
	创建关联数据表 t_movie_join, 并插入数据	30 分	能够正确创建数据表, 并插入数据	
	查询各电影的平均评分	20 分	能正确进行数据统计	
职业素养 (10分)	工作前准备	5 分	做好工作前准备, 检查电脑硬件 (键盘、鼠标等), 检查测试所需系统环境。不进行检查操作扣 5 分。	5 分
	职业行为规范	5 分	着装干净、整洁。举止文明, 遵守考场纪律, 按顺序进出考场。考试过程中及结束后, 考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	
总计		100 分		

项目 2: 网约车定位信息统计

2.试题编号: 3-2-1 《网约车定位信息统计》综合应用

(1) 任务描述

随着移动互联网的快速发展以及经济增长,网约车逐渐成为了国民出行的一种习惯,用户规模将稳定增长。对日益激增的定位信息需要更快速和高效地捕捉到有价值的信息。

现有网约车订单信息数据,存储在/root/experiment/datas/online_car.txt文件中,字段包括:订单时间(run_date),接单位置(start_site),目的位置(end_site)请使用 HiveQL 语句,完成以下功能。

- 1.创建数据表 t_online_car 并加载网约车订单信息数据。
- 2.统计接单位置订单数量。

任务一:启动 Hive 操作窗口(10分)

任务二:根据字段描述创建数据表 t_online_car(20分)

任务三:加载网约车订单信息数据到 Hive 表 t_online_car(40分)

任务四:查询各接单位置产生的订单数量(30分)

(2) 实施条件

抽测所需的软硬件如表 H1-4-1:

表 H1-4-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机(CPU2.0GHZ 以上,内存 16GB,硬盘空间 500G 以上);该计算机上搭载了或者能够访问 Hadoop 平台,该 Hadoop 平台上已安装好 Hive 数仓。	每人一台
	FTP 服务器 1 台考场服务器(CPU3.0GHZ 以上,内存 32GB 以上,硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	远程访问工具, Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家:在本行业具有 3 年以上从业经验(工程师及以上职称)或从事本专业或相关专业具有 5 年以上的教学经验(副高及以上职称)	测评专家满足任一条件
	结果测评专家:在本行业具有 3 年以上从业经验(工程师及以上职称)或从事本专业或相关专业具有 5 年以上的教学经验(副高及以上职称)。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制,评价内容包括工作任务、职业素养完成情况两个方面。其中,工作任务完成质量占该项目总分的 90%,职业素养占该项目总分的 10%。具体评价标准见下表:

表 H1-4-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注
工作任务 (90分)	启动 Hive 操作窗口	10 分	能够正确启动 hadoop 服务器,能启动 Hive	1、考试舞弊、抄袭、没有按要求
	根据字段描述创建数据表 t_online_car	15 分	能够正确创建数据表	

评价内容		配分	评价要点	备注
	加载网约车订单信息数据到 Hive 表 t_online_car	35 分	能够正确加载文件到数据表	填写相关信息，本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。
	查询各接单位位置产生的订单数量	30 分	能够查看表数据	
职业素养 (10 分)	工作前准备	5 分	做好工作前准备，检查电脑硬件（键盘、鼠标等），检查测试所需系统环境。不进行检查操作扣 5 分。	5 分
	职业行为规范	5 分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	
总计		100 分		

项目 3：酒店客服出租率统计

3.试题编号：3-3-1 《酒店客服出租率统计》综合应用

(1) 任务描述

随着我国经济的迅速发展，各地把旅游业当做本地经济发展的重要经济支柱之一，从而促进了酒店业的快速发展。通过酒店出租率统计分析，可以指导酒店管理者对整体运营进行科学管理。

现有酒店订单信息数据，存储在/root/experiment/datas/hotel_msg.txt 文件中，字段包括：订单日期（order_date），酒店名称（hotel_name），出租房间数量（out_num），总房间数量（all_num）

请使用 HiveQL 语句，完成以下功能。

- 1.创建数据表 t_hotel_msg 并加载网约车订单信息数据。
- 2.统计各酒店客房出租率。

任务一：启动 Hive 操作窗口（10 分）

任务二：根据字段描述创建数据表 t_hotel_msg（20 分）

任务三：加载数据到数据表（35 分）

任务四：查询查询各酒店客服出租率（35 分）

(2) 实施条件

抽测所需的软硬件如表 H1-4-1：

表 H1-4-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机（CPU2.0GHZ 以上，内存 16GB，硬盘空间 500G 以上）；该计算机上搭载了或者能够访问 Hadoop 平台，该 Hadoop 平台上已安装好 Hive 数仓。	每人一台
	FTP 服务器 1 台考场服务器（CPU3.0GHZ 以上，内存 32GB 以上，硬盘空间 1T 以上）；	用于保存测试人员考试结果

项目	基本实施条件	备注
工具	远程访问工具, Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制, 评价内容包括工作任务、职业素养完成情况两个方面。其中, 工作任务完成质量占该项目总分的 90%, 职业素养占该项目总分的 10%。具体评价标准见下表:

表 H1-4-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注
工作任务 (90分)	启动 Hive 操作窗口	10 分	能够正确启动 hadoop 服务器, 能启动 Hive	1、考试舞弊、抄袭、没有按要求填写相关信息, 本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。
	根据字段描述创建数据表 t_hotel_msg	20 分	能够正确创建数据表	
	加载数据到数据表	30	能够正确加载文件到数据表	
	查询查询各酒店客房出租率	30	能够查看表数据	
职业素养 (10分)	工作前准备	5 分	做好工作前准备, 检查电脑硬件 (键盘、鼠标等), 检查测试所需系统环境。不进行检查操作扣 5 分。	5 分
	职业行为规范	5 分	着装干净、整洁。举止文明, 遵守考场纪律, 按顺序进出考场。考试过程中及结束后, 考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	
总计		100 分		

项目 4: 薪水支出统计

4.试题编号: 3-4-1 《薪水支出统计》统计汇总实现

(1) 任务描述

现有用户薪酬信息数据存储在 /root/experiment/datas/ 下的 userinfo1.txt 文件中, 数据字段包括: 用户名 (user_name)、薪水 (user_salary)

请你用 HiveQL 语句加载数据, 并统计总薪水支出。

任务一: 启动 Hive 操作窗口

任务二: 根据数据字段描述在 Hive 默认数据库创建 t_userinfo 表

任务三：将 userinfo1.txt 加载到 t_userinfo 表

任务四：查看 t_userinfo 表数据

任务五：统计 t_userinfo 表薪水总和

(2) 实施条件

抽测所需的软硬件如表 H1-4-1:

表 H1-4-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 16GB, 硬盘空间 500G 以上); 该计算机上搭载了或者能够访问 Hadoop 平台, 该 Hadoop 平台上已安装好 Hive 数仓。	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	远程访问工具, Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制, 评价内容包括工作任务、职业素养完成情况两个方面。其中, 工作任务完成质量占该项目总分的 90%, 职业素养占该项目总分的 10%。具体评价标准见下表:

表 H1-4-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注
工作任务 (90分)	启动 Hive 操作窗口	10 分	能够正确启动 hadoop 服务器, 能启动 Hive	1、考试舞弊、抄袭、没有按 要求填写相关信息, 本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的 本项记 0 分。
	创建数据表	10 分	能够正确创建数据表	
	userinfo1.txt 加载到 t_userinfo 表	25	能够正确加载文件到数据表	
	查看 t_userinfo 表数据	15	能够查看表数据	
	统计 t_userinfo 表薪水总和	30	能够进行数据统计	
职业素养 (10分)	工作前准备	5 分	做好工作前准备, 检查电脑硬件 (键盘、鼠标等), 检查测试所需系统环境。不进行检查操作扣 5 分。	5 分
	职业行为规范	5 分	着装干净、整洁。举止文明,	5 分

评价内容	配分	评价要点	备注
		遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	
总计	100 分		

项目 5：股票日线数据清洗

5.试题编号：3-5-1 《股票日线数据清洗》数据清洗

(1) 任务描述

现有股票日线数据文件/root/experiment/datas/stocks.csv，包含字段：股票代码（sticks_id）、股票交易日期（stocks_date）、股票开盘价格（open_price）、股票收盘价格（close_price）。

请你用 HiveQL 语句进行数据清洗。

1. 根据数据字段描述创建 t_stocks 与 t_stocks_clear 数据表。
2. 去除股票交易日期为空的数据。

任务一：启动 Hive 操作窗口

任务二：创建数据表 t_stocks 用于存储原始数据

任务三：创建数据表 t_stocks_clear 用于存储清洗后数据

任务四：将 stocks.csv 数据加载到 t_stocks 数据表

任务五：编写 HiveQL 语句，查询 t_stocks 表中交易日期不为空的数据并插入到 t_stocks_clear 表

任务六：查询 t_stocks_clear 表中数据

(2) 实施条件

抽测所需的软硬件如表 H1-5-1：

表 H1-5-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机（CPU2.0GHZ 以上，内存 16GB，硬盘空间 500G 以上）；该计算机上搭载了或者能够访问 Hadoop 平台，该 Hadoop 平台上已安装好 Hive 数仓。	每人一台
	FTP 服务器 1 台考场服务器（CPU3.0GHZ 以上，内存 32GB 以上，硬盘空间 1T 以上）；	用于保存测试人员考试结果
工具	远程访问工具，Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家：在本行业具有 3 年以上从业经验（工程师及以上职称）或从事本专业或相关专业具有 5 年以上的教学经验（副高及以上职称）	测评专家满足任一条件
	结果测评专家：在本行业具有 3 年以上从业经验（工程师及以上职称）或从事本专业或相关专业具有 5 年以上的教学经验（副高及以上职称）。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 90%，职业素养占该项目总分的 10%。具体评价标准见下表：

表 H1-5-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注
工作任务 (90分)	启动 Hive 操作窗口	10 分	能够正确启动 hadoop 服务器，能启动 Hive	1、考试舞弊、抄袭、没有按要求填写相关信息，本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。
	创建数据表 t_stocks	15	能够正确创建数据表	
	创建数据表 t_stocks_clear 用于存储清洗后数据	15	能够正确创建数据表	
	将 stocks.csv 数据加载到 t_stocks 数据表	15	能够正确进行数据加载	
	编写 HiveQL 语句，查询 t_stocks 表中交易日期不为空的数据并插入到 t_stocks_clear 表	25	编写 HiveQL 语句，并将结果插入到数据表	
	查询 t_stocks_clear 表中数据	10	能够正确展示数据表中的数据	
职业素养 (10分)	工作前准备	5 分	做好工作前准备，检查电脑硬件（键盘、鼠标等），检查测试所需系统环境。不进行检查操作扣 5 分。	5 分
	职业行为规范	5 分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5 分
总计		100 分		

项目 6：学生消费数据清洗

6.试题编号：3-6-1 《学生消费数据清洗》去重操作实现

(1) 任务描述

现有学生校内消费数据，字段包括：学生 ID (std_id)，学生姓名 (std_name)，学生性别 (std_sex)，消费日期 (sale_date)，消费类型 (sale_type)，消费金额 (sale_num)。

数据存放在本地文件/root/experiment/datas/students.txt 中，请你用 HiveQL 语

句提取有消费记录的学生 ID、学生姓名、学生性别。

任务一：开启 Hive 操作窗口

任务二：创建学生消费信息表 t_students 并加载数据

任务三：创建去重后数据存储表 t_students_distinct

任务四：查询 Hive 中所有数据表

任务五：去重查询 t_students 中学生 ID、学生姓名、学生性别字段，并插入到 t_students_distinct 表中

任务六：查询 t_students_distinct 表数据内容

(2) 实施条件

抽测所需的软硬件如表 H1-6-1：

表 H1-6-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 16GB, 硬盘空间 500G 以上); 该计算机上搭载了或者能够访问 Hadoop 平台, 该 Hadoop 平台上已安装好 Hive 数仓。	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	远程访问工具, Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制, 评价内容包括工作任务、职业素养完成情况两个方面。其中, 工作任务完成质量占该项目总分的 90%, 职业素养占该项目总分的 10%。具体评价标准见下表:

表 H1-6-2 数据应用开发模块考核评分细则

评价内容	配分	评价要点	备注	
工作任务 (90 分)	启动 Hive 操作窗口	10 分	能够正确启动 hadoop 服务器, 能启动 Hive	1、考试舞弊、抄袭、没有按要求填写相关信息, 本模块记 0 分。 2、严重违反考场纪律、
	创建学生消费信息表 t_students 并加载数据	15	能够正确创建数据表	
	创建去重后数据存储表 t_students_distinct	15	能够正确创建数据表	
	查询 Hive 中所有数据表	15	能够进行正常查询	

评价内容		配分	评价要点	备注
	去重查询 t_students 中学生 ID、学生姓名、学生性别字段, 并插入到 t_students_distinct 表中	25	能够进行正确处理, 并能插入到数据表	造成恶劣影响的本项目记 0 分。
	查询 t_students_distinct 表数据内容	10	整个正确查询	
职业素养 (10 分)	工作前准备	5 分	做好工作前准备, 检查电脑硬件 (键盘、鼠标等), 检查测试所需系统环境。不进行检查操作扣 5 分。	5 分
	职业行为规范	5 分	着装干净、整洁。举止文明, 遵守考场纪律, 按顺序进出考场。考试过程中及结束后, 考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5 分
总计		100 分		

项目 7: 会员状况统计分析

7.试题编号: 3-7-1 《会员状况统计分析》综合应用

(1) 任务描述

会员营销是当下主流的营销手段之一, 通过会员管理掌握消费动态, 为企业决策提供助力。现有会员基础信息数据文件/root/experiment/datas/user_msg.txt。字段包括: 用户 ID (u_id), 用户性别 (u_sex), 用户名称 (u_name), 使用终端品牌 (手机品牌) (phone_type)。

请使用 HiveQL 语句, 完成以下功能:

- 1、统计会员总人数。
- 2、统计不同性别的会员人数。
- 3、统计不同终端的使用人数。

任务一: 开启 Hive 操作终端窗口

任务二: 根据数据字段创建数据表 t_user_msg, 并加载数据

任务三: 统计 t_user_msg 中用户总数

任务四: 统计不同性别的会员人数

任务五: 统计不同终端的使用人数

(2) 实施条件

抽测所需的软硬件如表 H1-7-1:

表 H1-7-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 16GB, 硬盘空间 500G 以上); 该计算机上搭载了或者能够访问 Hadoop 平台, 该 Hadoop 平台上已安装好 Hive 数仓。	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB)	用于保存测试人

项目	基本实施条件	备注
	以上，硬盘空间 1T 以上);	员考试结果
工具	远程访问工具，Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家：在本行业具有 3 年以上从业经验（工程师及以上职称）或从事本专业或相关专业具有 5 年以上的教学经验（副高及以上职称）	测评专家满足任一条件
	结果测评专家：在本行业具有 3 年以上从业经验（工程师及以上职称）或从事本专业或相关专业具有 5 年以上的教学经验（副高及以上职称）。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 90%，职业素养占该项目总分的 10%。具体评价标准见下表：

表 H1-7-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注
工作任务 (90分)	启动 Hive 操作窗口	10 分	能够正确启动 hadoop 服务器，能启动 Hive	1、考试舞弊、抄袭、没有按要求填写相关信息，本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。
	根据数据字段创建数据表 t_user_msg，并加载数据	10	能够进行正确创建和加载	
	统计 t_user_msg 中用户总数	20	能够进行正确统计	
	统计不同性别的会员人数	20	能够进行正确统计	
	统计不同终端的使用人数	30	能够进行正确统计	
职业素养 (10分)	工作前准备	5 分	做好工作前准备，检查电脑硬件（键盘、鼠标等），检查测试所需系统环境。不进行检查操作扣 5 分。	5 分
	职业行为规范	5 分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	
总计		100 分		

项目 8：金融风控分析

8.试题编号：3-8-1 《金融风控分析》综合应用

(1) 任务描述

金融风控是金融交易过程中的一个重要环节，是核心。风控不仅关系经营的

成本高低，更重要的是直接影响到投资人的资金安全。现有部分金融风控数据，存储在/root/experiment/datas/fengkong.txt 文件中，字段内容包括：用户 ID(u_id)，审核申请时间(examine_time)，认证间隔时间(interval_time)，订单总额(order_money)，审核标识(1 通过，0 未通过)(examine_flag)

请使用 HiveQL 语句提取通过审核的业务数据，并统计通过审核数据中订单总额的平均值。

任务一：开启 Hive 操作终端窗口

任务二：根据数据字段创建数据表 t_fengkong，并加载数据

任务三：创建数据表 t_fengkong_clear，用于存储通过审核的业务数据

任务四：查询 t_fengkong 表中通过审核的数据，并插入到 t_fengkong_clear 表中

任务五：统计 t_fengkong_clear 表中订单总额的平均值

(2) 实施条件

抽测所需的软硬件如表 H1-7-1:

表 H1-7-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上，内存 16GB，硬盘空间 500G 以上)；该计算机上搭载了或者能够访问 Hadoop 平台，该 Hadoop 平台上已安装好 Hive 数仓。	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上，内存 32GB 以上，硬盘空间 1T 以上)；	用于保存测试人员考试结果
工具	远程访问工具，Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家：在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)	测评专家满足任一条件
	结果测评专家：在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 90%，职业素养占该项目总分的 10%。具体评价标准见下表：

表 H1-7-2 数据应用开发模块考核评分细则

评价内容	配分	评价要点	备注	
工作任务 (90 分)	启动 Hive 操作窗口	10 分	能够正确启动 hadoop 服务器，能启动 Hive	1、考试舞弊、抄袭、没有按要求填写相关信息，本模块
	根据数据字段创建数据表 t_fengkong，并加载数据	10	能够进行正确创建和加载	
	创建数据表 t_fengkong_clear	20	能够进行正确创建数据表	

评价内容		配分	评价要点	备注	
	查询 t_fengkong 表中通过审核的数据，并插入到 t_fengkong_clear 表中	30	能够进行正确数据插入	记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。	
	统计 t_fengkong_clear 表中订单总额的平均值	20	能够进行正确统计		
职业素养 (10分)	工作前准备	5分	做好工作前准备，检查电脑硬件（键盘、鼠标等），检查测试所需系统环境。不进行检查操作扣 5 分。		5分
	职业行为规范	5分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。		5分
总计		100分			

项目 9：店铺零售业绩统计分析

9.试题编号：3-9-1 《店铺零售业绩统计分析》综合应用

(1) 任务描述

店铺零售行业是所有行业中对大数据最为敏感的行业之一。想要赚取最大利润往往需要对消费者的购买趋势做出最快反应，使用大数据来分析消费趋势以此在抢购热潮中获得更多收益。

现有消费明细数据与商品信息数据。分别存储在

/root/experiment/datas/sale.txt、/root/experiment/datas/goods.txt 文件中。

消费明细数据包含字段：销售日期 (sale_date)，会员 ID (user_id)，销售商品 ID (goods_id)

商品信息数据包含字段：商品 ID (goods_id)，商品名称 (goods_name)，销售价格 (goods_price)

请使用 HiveQL 语句，完成以下功能。

- 1.以商品 ID 作为关联字段，关联消费明细数据与商品信息数据
- 2.根据关联数据表统计每种商品的销售总额。

任务一：开启 Hive 操作终端窗口

任务二：根据字段描述创建销售明细表 t_sale，并加载数据

任务三：根据字段描述创建商品信息表 t_goods，并加载数据

任务四：创建关联数据表 t_join，字段包括销售日期，会员 ID，销售商品 ID，商品名称，销售价格

任务五：以商品 ID 作为关联条件关联 t_sale 与 t_goods，并将关联数据插入 t_join

任务六：统计 t_join 表中每种商品的销售总额，查询字段包括商品 ID、商

品名称、销售总额，并以商品 ID 进行排序

(2) 实施条件

抽测所需的软硬件如表 H1-9-1:

表 H1-9-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 16GB, 硬盘空间 500G 以上); 该计算机上搭载了或者能够访问 Hadoop 平台, 该 Hadoop 平台上已安装好 Hive 数仓。	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	远程访问工具, Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制, 评价内容包括工作任务、职业素养完成情况两个方面。其中, 工作任务完成质量占该项目总分的 90%, 职业素养占该项目总分的 10%。具体评价标准见下表:

表 H1-9-2 数据应用开发模块考核评分细则

评价内容	配分	评价要点	备注	
工作任务 (90 分)	启动 Hive 操作窗口	10	能够正确启动 hadoop 服务器, 能启动 Hive	1、考试舞弊、抄袭、没有按要求填写相关信息, 本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。
	创建销售明细表 t_sale, 并加载数据	10	能够进行正确创建数据表和数据加载	
	创建商品信息表 t_goods, 并加载数据	10	能够进行正确创建数据表和数据加载	
	创建关联数据表 t_join	20	能够进行正确进行数据关联	
	以商品 ID 作为关联条件关联 t_sale 与 t_goods, 并将关联数据插入 t_join	20	能够进行数据插入	
	统计 t_join 表中每种商品的销售总额	20	能够进行正确统计	
职业素养 (10)	工作前准备	5 分	做好工作前准备, 检查电脑硬件 (键盘、鼠标等), 检查测试所需系统环境。不进行检查	5 分

评价内容	配分	评价要点	备注
分)		操作扣 5 分。	
职业行为规范	5 分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5 分
总计	100 分		

项目 10：面积统计分析

10.试题编号：3-10-1 《面积统计分析》综合应用

(1) 任务描述

农业耕地面积的分析是我国农业发展分析的重要组成部分,对于发展农业现代化具有重要的参考意义。现有农业种植面积数据,存储在 /root/experiment/datas/farming.txt 文件中,字段包括:年份(date_year),省份(shen_name),行政区(qu_name),农作物类型(farm_type),种植面积(farm_area)。

请使用 HiveQL 语句,完成以下功能。

- 1.统计种植面积前 5 的省份。
- 2.统计种植面积前 5 的农作物类型。

任务一：开启 Hive 操作终端窗口

任务二：根据字段描述创建销售明细表 t_farming, 并加载数据

任务三：创建省份种植面积统计表 t_area 与农作物类型种植面积统计表 t_crops

任务四：根据 t_farming 表中数据按省份统计总种植面积, 并插入 t_area 数据表

任务五：查询种植面积前 5 的省份名称与总种植面积

任务六：根据 t_farming 表中数据按农作物类型统计总种植面积, 并插入 t_crops 数据表

任务七：查询种植面积前 5 的农作物类型与总种植面积

(2) 实施条件

抽测所需的软硬件如表 H1-10-1:

表 H1-10-1 大数据 Hadoop 综合模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 16GB, 硬盘空间 500G 以上); 该计算机上搭载了或者能够访问 Hadoop 平台, 该 Hadoop 平台上已安装好 Hive 数仓。	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	远程访问工具, Xshell 或者其他工作	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称)	测评专家满足任一条件

项目	基本实施条件	备注
	结果测评专家：在本行业具有3年以上从业经验（工程师及以上职称）或从事本专业或相关专业具有5年以上的教学经验（副高及以上职称）。	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

大数据 Hadoop 综合模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 90%，职业素养占该项目总分的 10%。具体评价标准见下表：

表 H1-10-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注
工作任务 (90分)	启动 Hive 操作窗口	5	能够正确启动 hadoop 服务器，能启动 Hive	1、考试舞弊、抄袭、没有按要求填写相关信息，本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。
	创建销售明细表 t_farming，并加载数据	10	能够正确创建数据表，并进行数据加载	
	创建省份种植面积统计表 t_area 与农作物类型种植面积统计表 t_crops	15	能够正确创建数据表	
	根据 t_farming 表中数据按省份统计总种植面积，并插入 t_area 数据表	15	能够正确统计并进行数据插入	
	查询种植面积前 5 的省份名称与总种植面积	15	能够进行正确查询	
	根据 t_farming 表中数据按农作物类型统计总种植面积，并插入 t_crops 数据表	15	能够正确统计并进行数据插入	
	查询种植面积前 5 的农作物类型与总种植面积	15	能够进行统计查询	
职业素养 (10分)	工作前准备	5分	做好工作前准备，检查电脑硬件（键盘、鼠标等），检查测试所需系统环境。不进行检查操作扣 5 分。	5分
	职业行为规范	5分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5分
总计		100分		

模块四 大数据技术应用

项目 1: 北京空气质量监测数据获取与分析

1. 试题编号: 4-1-1 北京空气质量监测数据获取与分析

(1) 任务描述

任务 1: 数据采集

网页“<http://pm25.in/beijing>”中包含北京 12 个监测点的空气质量监测数据, 请编写程序抓取网页(网页样本保存在源素材文件夹下 `src1` 目录中)上的监测点、AQI、空气质量指数类别, 如表 H2-1-1 所示。将抓取的数据保存, 以 `bj20200721.csv` 文件命名。

表 H2-1-1 北京空气质量监测数据

监测点	AQI	空气质量指数类别
万寿西宫	57	良
.....

1) 使用 Pycharm 在考生文件夹下创建项目 `task030101`, 在项目下创建 python 文件, 命名为 `task1.py`。将源素材文件夹下的 `src1` 目录复制到项目 `task030101` 下。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 正确爬取网页中的数据, 运行程序并显示“已爬取”。

4) 编写程序, 将抓取的数据保存在项目 `task030101` 下, 以 `bj20200721.csv` 文件命名, 分隔符为 `tab` 键, 运行程序并显示“已保存”。

5) 将运行结果截图, 保存在考生文件夹下, 以 `task1.jpg` 文件命名。

任务 2: 数据处理和分析

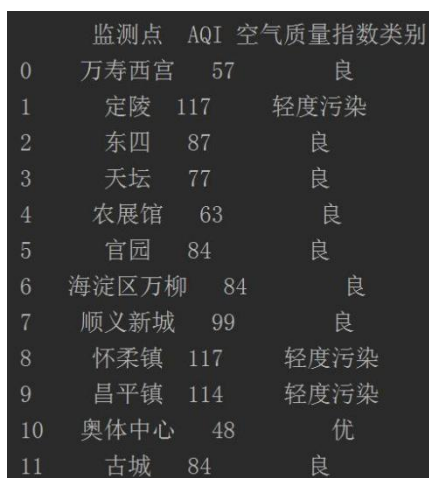
使用 Python 的 `pandas` 读取 `bj20200721.csv` 文件(`bj20200721.csv` 文件保存在源素材文件夹下 `task030102\src2` 目录中), 完成数据处理和分析。

1) 将源素材文件夹下 `task030102` 项目复制到考生文件夹下, 使用 Pycharm 打开项目 `task030102`, 打开 `task2.py`。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 完成 `read_file(filename)` 函数的定义, 使用 `pandas` 读取 `bj20200721.csv` 文件中的数据, 返回值类型为 `DataFrame`。

4) 编写程序, 完成 `print_info(filename)` 函数定义, 显示 `bj20200721.csv` 文件中的内容, 如图 H2-1-1 所示。



	监测点	AQI	空气质量指数类别
0	万寿西宫	57	良
1	定陵	117	轻度污染
2	东四	87	良
3	天坛	77	良
4	农展馆	63	良
5	官园	84	良
6	海淀区万柳	84	良
7	顺义新城	99	良
8	怀柔镇	117	轻度污染
9	昌平镇	114	轻度污染
10	奥体中心	48	优
11	古城	84	良

图 H2-1-1 显示 `bj20200721.csv` 文件中的内容

5) 编写程序, 完成 `insert_date(filename)` 函数的定义, 将 2020-07-21 日期数据作为新的一列添加到 Excel 文件中, 列名称为“监测日期”, 保存在项目 `task030102` 下, 文件命名为 `bj20200721.xlsx`。

6) 编写程序, 完成 `aqi(filename)` 函数的定义, 显示空气质量指数类别为优的记录。

7) 编写程序, 导入模块, 调用自定义的 `read_file`、`print_info`、`insert_date`、`aqi` 函数, 实现上述功能, 注意参数的正确设置。

8) 将运行结果截图, 保存在考生文件夹下, 以 `task2.jpg` 文件命名。

任务 3: 数据可视化

使用 Python 读取 `bj20200721.csv` 文件 (`bj20200721.csv` 文件保存在源素材文件夹下 `src3` 目录中) 中的数据, 使用 `matplotlib` 绘制直方图, 显示 `aqi` 最少的 4 个监测点。

1) 使用 Pycharm 在考生文件夹下创建项目 `task030103`, 在项目下创建 python 文件, 命名为 `task3.py`。将源素材文件夹下的 `src3` 目录复制到项目 `task030103` 下。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 使用 `pandas` 读取 `bj20200721.csv` 文件中的数据。使用 `matplotlib` 绘制直方图, 显示 AQI 最少的 5 个监测点, 如图 H2-1-2 所示。生成的文件保存在项目 `task030103` 下, 文件命名为 `aqi.png`。

4) 将运行结果截图, 保存在考生文件夹下, 以 `task3.jpg` 文件命名。

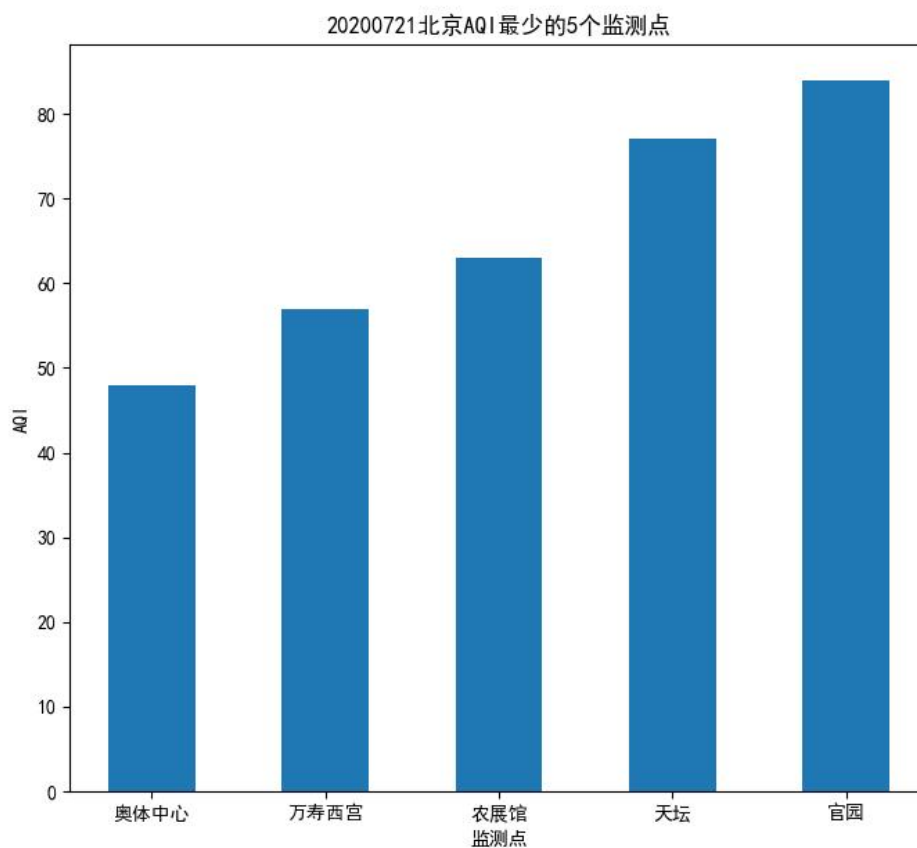


图 H2-1-2 AQI 最少的 5 个监测点直方图

(2) 实施条件

抽测所需的软硬件如表 H2-1-1:

表 H2-1-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	Microsoft Office 2010 以上版本; MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019 安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pyecharts1.x 等模块	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制, 评价内容包括工作任务、职业素养完成情况两个方面。其中, 工作任务完成质量占该项目总分的 80%, 职业素养占该项目总分的 20%。具体评价标准见下表:

表 H2-1-2 数据应用开发模块考核评分细则

评价内容	配分	评价要点	备注		
工作任务 (80 分)	数据 采集	新建、保存、运行项目和文件正确。	5 分	1、考试舞弊、抄袭、没有按要求填写相关信息, 本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0	
		认识 HTML 脚本语言, 网页抓包分析正确, 运行结果进行了截图。截图不正确扣 1 分。	10 分		
		能根据实际项目需求进行问题分析, 使用 Python 网络爬虫技术爬取网页中的数据。代码不正确 1 处扣 1 分。	10 分		
		将数据写入 csv 文件。代码不正确 1 处扣 1 分。	10 分		
	数据 处理和 析	25 分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图, 截图不正确扣 2 分。		5 分
			read_file(filename)函数定义正确。代码不正确 1 处扣 1 分。		5 分
			print_info(filename)函数定义正确。代码不正确 1 处扣 1 分。		5 分
			insert_date(filename)函数定义正确。代码不正确 1 处扣 1 分。		5 分

评价内容		配分	评价要点	备注
数据可视化	20分	5分	aqi(filename)函数定义正确。代码不正确1处扣1分。	备注分。
		5分	新建、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣1分。	
		2分	导入 pandas、matplotlib 模块正确	
		1分	中文显示正确。	
		2分	读取 csv 文件正确。	
		2分	设置标题正确。	
		2分	按 AQI 列升序排序正确。	
		2分	绘制直方图正确。	
		2分	X 轴标签和 Y 轴标签正确。	
职业素养 (20分)	5分	5分	做好工作前准备, 检查电脑硬件(键盘、鼠标等), 检查测试所需软件开发环境。不进行检查操作扣5分。	
	10分	10分	代码符合软件开发规范, 命名规范, 能做到见名知意; 缩进统一, 方便阅读; 注释规范正确, 至少要有5行注释, 少1行或错1行扣1分。	
	5分	5分	着装干净、整洁。举止文明, 遵守考场纪律, 按顺序进出考场。考试过程中及结束后, 考试桌面及地面不符合6S基本要求的扣3—5分。	
总计	100分			

项目 2: 58921 电影票房数据获取与分析

2. 试题编号: 4-2-1 58921 电影票房数据获取与分析

(1) 任务描述

任务 1: 数据采集

网页“<http://58921.com/boxoffice/wangpiaopiao/20190701>”中包含 2019 年 7 月 1 日网票票房相关统计信息, 请编写程序抓取网页(网页样本保存在源素材文件夹下 src1 目录中)上的电影、总票房(元)、废场、人次数据, 如表 H2-2-1 所示。将抓取的数据保存, 以 film20190701.csv 文件命名。

表 H2-2-1 电影票房数据

电影	总票房(元)	废场	人次
蜘蛛侠: 英雄远征	2.43 亿	83808	768745
.....

1) 使用 Pycharm 在考生文件夹下创建项目 task030201, 在项目下创建 python 文件, 命名为 task1.py。将源素材文件夹下的 src1 目录复制到项目 task030201 下。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 正确爬取网页中的数据, 运行程序并显示结果。

4) 编写程序, 将抓取的数据保存在项目 task030201 下, 以 film20190701.csv 文件命名, 分隔符为 tab 键。

5) 运行结果截图, 保存在考生文件夹下, 以 task1.jpg 文件命名。

任务 2: 数据处理和分析

使用 Python 的 pandas 读取 20190801.csv 文件(20190801.csv 文件保存在源

素材文件夹下 task030202\src2 目录中), 完成数据处理和分析。

1) 将源素材文件夹下 task030202 项目复制到考生文件夹下, 使用 Pycharm 打开项目 task030202, 打开 task2.py。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 完成 read_file(filename)函数的定义, 使用 pandas 读取 20190801.csv 文件中的数据, 返回值类型为 DataFrame。

4) 编写程序, 完成 print_info(filename)函数定义, 显示 20190801.csv 文件中前 5 行数据, 如图 H2-2-1 所示。



	电影	总票房 (元)	废场	人次
0	哪吒之魔童降世	3.87亿	76529	1399357
1	烈火英雄	5537.69万	63392	1114448
2	古田军号	493.91万	15858	133093
3	银河补习班	2.73亿	15121	90438
4	狮子王	2.77亿	4937	23965

图 H2-2-1 显示前 5 行数据

5) 编写程序, 完成 insert_date(filename)函数的定义, 将 2019-08-01 日期数据作为新的一列添加到 Excel 文件中, 列名为“上映日期”, 保存在项目 task030202 下, 文件命名为 20190801.xlsx。

6) 编写程序, 完成 max(filename)函数的定义, 显示人次最多的电影的信息。

7) 编写程序, 导入模块, 调用自定义的 read_file、print_info、insert_date、max 函数, 实现上述功能, 注意参数的正确设置。

8) 运行结果截图, 保存在考生文件夹下, 以 task2.jpg 文件命名。

任务 3: 数据可视化

使用 Python 读取 20190801.csv 文件 (20190801.csv 文件保存在源素材文件夹下 src3 目录) 中的数据, 使用 matplotlib 绘制折线图, 显示观看人次最高的 5 部电影。

1) 使用 Pycharm 在考生文件夹下创建项目 task030203, 在项目下创建 python 文件, 命名为 task3.py。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 使用 pandas 读取 20190801.csv 文件中的数据。使用 matplotlib 绘制折线图, 显示观看人次最低的 5 部电影, 如图 H2-2-2 所示。生成的文件保存在项目 task030203 下, 文件命名为 film.png。

4) 运行结果截图, 保存在考生文件夹下, 以 task3.jpg 文件命名。

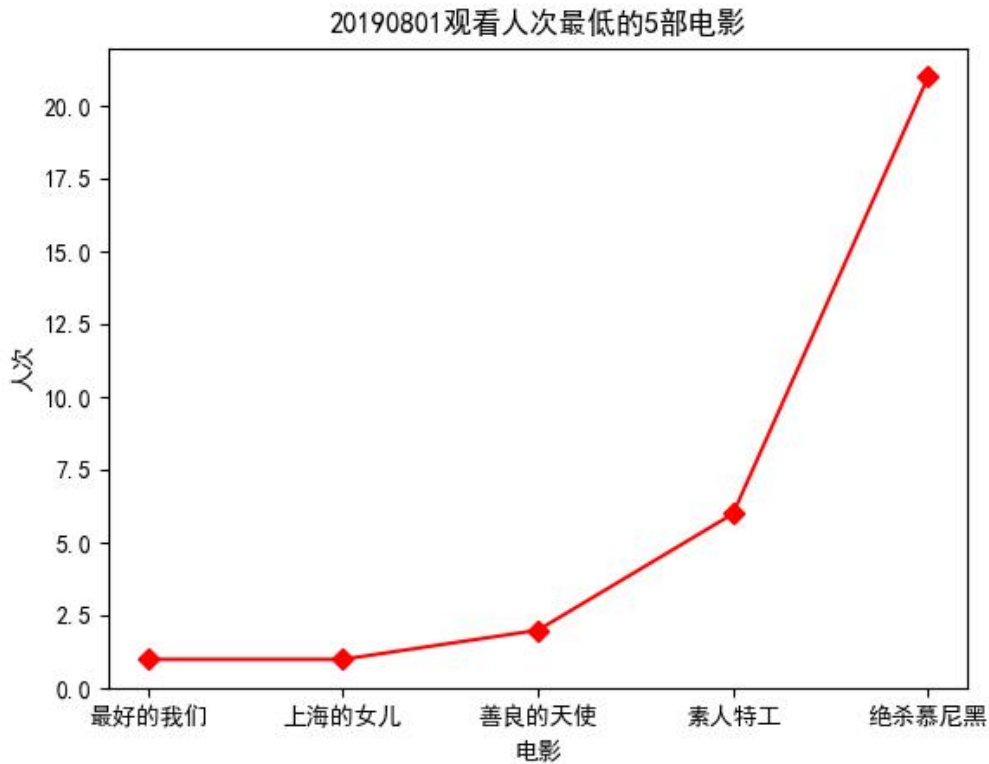


图 H2-2-2 观看人次最高的 5 部电影折线图

(2) 实施条件

抽测所需的软硬件如表 H2-2-1:

表 H2-2-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	Microsoft Office 2010 以上版本; MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019 安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pycharts1.x 等模块	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 80%，职业素养占该项目总分的 20%。具体评价标准见下表：

表 H2-2-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点		备注
工作任务 (80分)	数据采集	35分	新建、保存、运行项目和文件正确。	5分	1、考试舞弊、抄袭、没有按 要求填写相关信息， 本模块记0分。 2、严重违反考场纪 律、造成恶劣影响的 本项记0分。
			认识 HTML 脚本语言，网页抓包分析正确，运行结果进行了截图。截图不正确扣 1 分。	10分	
			能根据实际项目需求进行问题分析，使用 Python 网络爬虫技术爬取网页中的数据。代码不正确 1 处扣 1 分。	10分	
			将数据写入 csv 文件。代码不正确 1 处扣 1 分。	10分	
	数据处理和析	25分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 2 分。	5分	
			read_file(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			print_info(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			insert_date(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			max(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
	数据可视化	20分	新建、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 1 分。	5分	
			导入 pandas、matplotlib 模块正确	2分	
			中文显示正确。	1分	
			读取 csv 文件正确。	2分	
			设置标题正确。	2分	
			按人次列升序排序正确。	2分	
			绘制折线图正确。	2分	
			X 轴标签和 Y 轴标签正确。	2分	
	保存和显示图片正确。	2分			
职业素养 (20分)	工作前准备	5分	做好工作前准备，检查电脑硬件(键盘、鼠标等)，检查测试所需软件开发环境。不进行检查操作扣 5 分。	5分	
	专业素养	10分	代码符合软件开发规范，命名规范，能做到见名知意；缩进统一，方便阅读；注释规范正确，至少要有 5 行注释，少 1 行或错 1 行扣 1 分。	10分	
	职业行为规范	5分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5分	
	总计	100分			

项目 3：地表水水质监测数据获取与分析

3.试题编号：4-3-1 地表水水质监测数据获取与分析

(1) 任务描述

任务 1：数据采集

网页“<http://ditu.92cha.com/shuizhi.php>”中包含地表水水质监测数据，请编写程序抓取网页（网页样本保存在源素材文件夹下 src1 目录中）上的断面名称、测量时间、pH、水质类别、站点情况，如表 H2-3-1 所示。将抓取的数据保存，以 water.csv 文件命名。

表 H2-3-1 地表水水质监测数据

断面名称	测量时间	pH	溶解氧	氨氮
北京古北口	2018-09-17 16:05	7.7	10.38	0.19
.....

- 1) 使用 Pycharm 在考生文件夹下创建项目 task030301，在项目下创建 python 文件，命名为 task1.py。将源素材文件夹下的 src1 目录复制到项目 task030301 下。
- 2) 根据实际项目需求进行问题分析，书写注释。
- 3) 编写程序，正确爬取网页中的数据，运行程序并显示“已爬取”。
- 4) 编写程序，将抓取的数据保存在项目 task030301 下，以 water.csv 文件命名，运行程序并显示“已保存”。
- 5) 将运行结果截图，保存在考生文件夹下，以 task1.jpg 文件命名。

任务 2：数据处理和分析

使用 Python 的 pandas 读取 water.csv 文件（water.csv 文件保存在源素材文件夹下 task030302\src2 目录中），完成数据处理和分析。

- 1) 将源素材文件夹下 task030302 项目复制到考生文件夹下，使用 Pycharm 打开项目 task030302，打开 task2.py。
- 2) 根据实际项目需求进行问题分析，书写注释。
- 3) 编写程序，完成 read_file(filename)函数的定义，使用 pandas 读取 water.csv 文件中的数据，返回值类型为 DataFrame。
- 4) 编写程序，完成 print_info(filename)函数定义，显示 water.csv 文件中最后 5 行的内容，如图 H2-3-1 所示。

	断面名称	测量时间	pH	水质类别	站点情况
60	四川宜宾凉姜沟	2020-04-25 04:00	7.69	II	正常
61	云南红河州河口	2020-05-25 08:00	7.66	II	正常
62	陕西渭南潼关吊桥	2020-05-25 08:08	8.44	II	正常
63	甘肃兰州新城桥	2020-05-25 08:00	--	--	河流断流
64	宁夏中卫新墩	2020-05-25 12:00	8.04	I	仪器故障

图 H2-3-1 显示 water.csv 文件最后 5 行内容

- 5) 编写程序，完成 del_before_date(filename)函数的定义，将 2020-01-0100:00:00 之后监测的数据保存在项目 task030302 下，文件命名为 water.xlsx。
- 6) 编写程序，完成 water_type(filename)函数的定义，显示水质类别非‘--’的数据记录。
- 7) 编写程序，导入模块，调用自定义的 read_file、print_info、del_before_date、water_type 函数，实现上述功能，注意参数的正确设置。
- 8) 将运行结果截图，保存在考生文件夹下，以 task2.jpg 文件命名。

任务 3：数据可视化使用

Python 读取 water3.csv 文件（water3.csv 文件保存在源素材文件夹下 src3 目录中）中的数据，计算 I 类水、II 类水、III 类水、IV 类水、V 类水、劣 V 类水记录的数量，使用 matplotlib 绘制饼图，显示 I 类水、II 类水、III 类水、IV 类水、V 类水、劣 V 类水所占比例。

1) 使用 Pycharm 在考生文件夹下创建项目 task030303，在项目下创建 python 文件，命名为 task3.py。将源素材文件夹下的 src3 目录复制到项目 task030303 下。

2) 根据实际项目需求进行问题分析，书写注释。

3) 编写程序，使用 pandas 读取 water3.csv 文件中的数据，计算 I 类水、II 类水、III 类水、IV 类水、V 类水、劣 V 类水记录的数量。

4) 使用 matplotlib 绘制饼图，显示 I 类水、II 类水、III 类水、IV 类水、V 类水、劣 V 类水所占比例，如图 H2-3-2 所示。生成的文件保存在项目 task030303 下，文件命名为 water.png。

5) 将运行结果截图，保存在考生文件夹下，以 task3.jpg 文件命名。

水质类别比例图

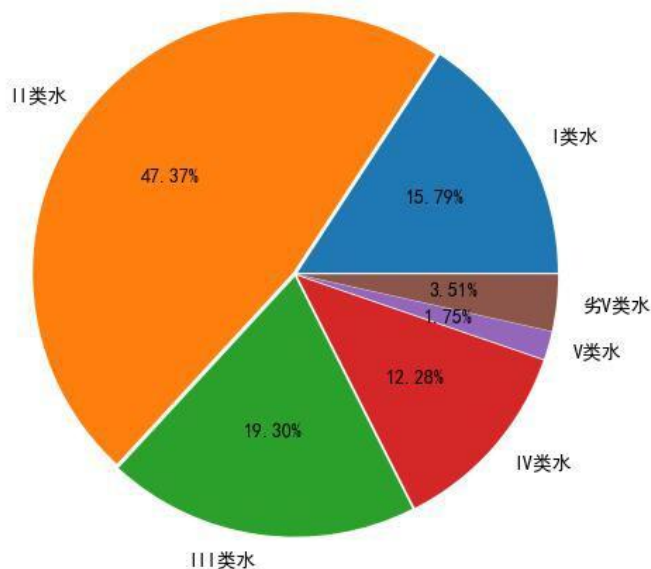


图 H2-3-2 水质类别饼图

(2) 实施条件

抽测所需的软硬件如表 H2-3-1：

表 H2-3-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	Microsoft Office 2010 以上版本; MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019	参考人员使用软件进行工作

项目	基本实施条件	备注
	安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pyecharts1.x 等模块	
测评专家	现场测评专家：在本行业具有 3 年以上从业经验（工程师及以上职称）或从事本专业或相关专业具有 5 年以上的教学经验（副高及以上职称），或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书（2 人/场）	测评专家满足任一条件
	结果测评专家：在本行业具有 3 年以上从业经验（工程师及以上职称）或从事本专业或相关专业具有 5 年以上的教学经验（副高及以上职称），或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书（2 人/场）	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 80%，职业素养占该项目总分的 20%。具体评价标准见下表：

表 H2-3-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注	
工作任务 (80分)	数据采集	35分	新建、保存、运行项目和文件正确。	5分	1、考试舞弊、抄袭、没有按要求填写相关信息，本模块记0分。 2、严重违反考场纪律、造成恶劣影响的本项目记0分。
			认识 HTML 脚本语言，网页抓包分析正确，运行结果进行了截图。截图不正确扣 1 分。	10分	
			能根据实际项目需求进行问题分析，使用 Python 网络爬虫技术爬取网页中的数据。代码不正确 1 处扣 1 分。	10分	
			将数据写入 csv 文件。代码不正确 1 处扣 1 分。	10分	
	数据分析和处理	25分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 2 分。	5分	
			read_file(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			print_info(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			del_before_date(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			water_type(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
	数据可视化	20分	新建、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 1 分。	5分	
			导入 pandas、matplotlib 模块正确	2分	
			中文显示正确。	1分	
			读取 csv 文件正确。	2分	
			设置标题正确。	2分	
			计算 I 类水、II 类水、III 类水、IV 类水、V 类水、劣 V 类水记录的数量正确。	2分	
			绘制饼图正确。	2分	
	X 轴标签和 Y 轴标签正确。	2分			
	保存和显示图片正确。	2分			

评价内容		配分	评价要点	备注
职业素养 (20分)	工作准备	5分	做好工作前准备,检查电脑硬件(键盘、鼠标等),检查测试所需软件开发环境。不进行检查操作扣5分。	5分
	专业素养	10分	代码符合软件开发规范,命名规范,能做到见名知意;缩进统一,方便阅读;注释规范正确,至少要有5行注释,少1行或错1行扣1分。	10分
	职业行为规范	5分	着装干净、整洁。举止文明,遵守考场纪律,按顺序进出考场。考试过程中及结束后,考试桌面及地面不符合6S基本要求的扣3—5分。	5分
总计		100分		

项目 4: 空气吸收剂量率数据获取与分析

4.试题编号: 4-4-1 空气吸收剂量率数据获取与分析

(1) 任务描述

任务 1: 数据采集

网页“<http://ditu.92cha.com/kqxsjll.php>”中包含空气吸收剂量率实时数据,请编写程序抓取网页(网页样本保存在源素材文件夹下 src1 目录中)上的省份、监测点、测量时间、测量值(nGy/h),如表 H2-4-1 所示。将抓取的数据保存,以 kqxsjll.csv 文件命名。

表 H2-4-1 空气吸收剂量率数据

省份	监测点	测量时间	测量值(nGy/h)
北京	奥林匹克森林公园站	2020-07-2701:30	78.2
.....

1)使用 Pycharm 在考生文件夹下创建项目 task030401,在项目下创建 python 文件,命名为 task1.py。将源素材文件夹下的 src1 目录复制到项目 task030401 下。

2)根据实际项目需求进行问题分析,书写注释。

3)编写程序,正确爬取网页中的数据,运行程序并显示“已爬取”。

4)编写程序,将抓取的数据保存在项目 task030401 下,以 kqxsjll.csv 文件命名,运行程序并显示“已保存”。

5)将运行结果截图,保存在考生文件夹下,以 task1.jpg 文件命名。

任务 2: 数据处理和分析

使用 Python 的 pandas 读取 kqxsjll.csv 文件(kqxsjll.csv 文件保存在源素材文件夹下 task030402\src2 目录中),完成数据处理和分析。

1)将源素材文件夹下 task030402 项目复制到考生文件夹下,使用 Pycharm 打开项目 task030402,打开 task2.py。

2)根据实际项目需求进行问题分析,书写注释。

3)编写程序,完成 read_file(filename)函数的定义,使用 pandas 读取 kqxsjll.csv 文件中的数据,返回值类型为 DataFrame。

4)编写程序,完成 print_info(filename)函数定义,显示 kqxsjll.csv 文件中前 10 行的内容,如图 H2-4-1 所示。

	省份	监测点	测量时间	测量值 (nGy/h)
0	北京	奥林匹克森林公园站	2020-07-27 01:30	78.2
1	北京	北京万柳中路站	2020-07-27 01:30	89.1
2	北京	北京门头沟西苑路站	2020-07-27 01:30	66.1
3	北京	北京昌平昌赤路站	2020-07-27 01:30	71.1
4	北京	亦庄博大公园站	2020-07-27 01:30	73.5
5	北京	北京平谷赵赵路站	2020-07-27 01:30	65.0
6	北京	北京延庆夏都公园站	2020-07-27 01:30	70.2
7	天津	南开复康路站	2020-07-27 01:30	72.3
8	天津	天津津南咸水沽小学站	2020-07-27 01:30	69.9
9	天津	天津北辰西平道站	2020-07-27 01:30	60.2

图 H2-4-1 显示 kqxsjll.csv 文件前 10 行内容

5) 编写程序, 完成 `hunan(filename)` 函数的定义, 将湖南监测点的数据保存在项目 `task030402` 下, 文件命名为 `hunan.xlsx`。

6) 编写程序, 完成 `sort(filename)` 函数的定义, 按测量值 (nGy/h) 从小到大排序, 显示运行结果。

7) 编写程序, 导入模块, 调用自定义的 `read_file`、`print_info`、`hunan`、`sort` 函数, 实现上述功能, 注意参数的正确设置。

8) 将运行结果截图, 保存在考生文件夹下, 以 `task2.jpg` 文件命名。

任务 3: 数据可视化

使用 Python 读取 `data.csv` 文件 (`data.csv` 文件保存在源素材文件夹下 `src3` 目录中) 中的数据, 使用 `matplotlib` 绘制散点图, 显示所有监测点的测量值 (nGy/h) 分布情况。

1) 使用 Pycharm 在考生文件夹下创建项目 `task030403`, 在项目下创建 python 文件, 命名为 `task3.py`。将源素材文件夹下的 `src3` 目录复制到项目 `task030403` 下。

2) 根据实际项目需求进行问题分析, 书写注释。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 使用 `pandas` 读取 `data.csv` 文件中的数据。使用 `matplotlib` 绘制折线图, 显示测量值 (nGy/h) 最低的 5 个测量点, 如图 H2-2-2 所示。生成的文件保存在项目 `task030203` 下, 文件命名为 `data.png`。

4) 运行结果截图, 保存在考生文件夹下, 以 `task3.jpg` 文件命名。

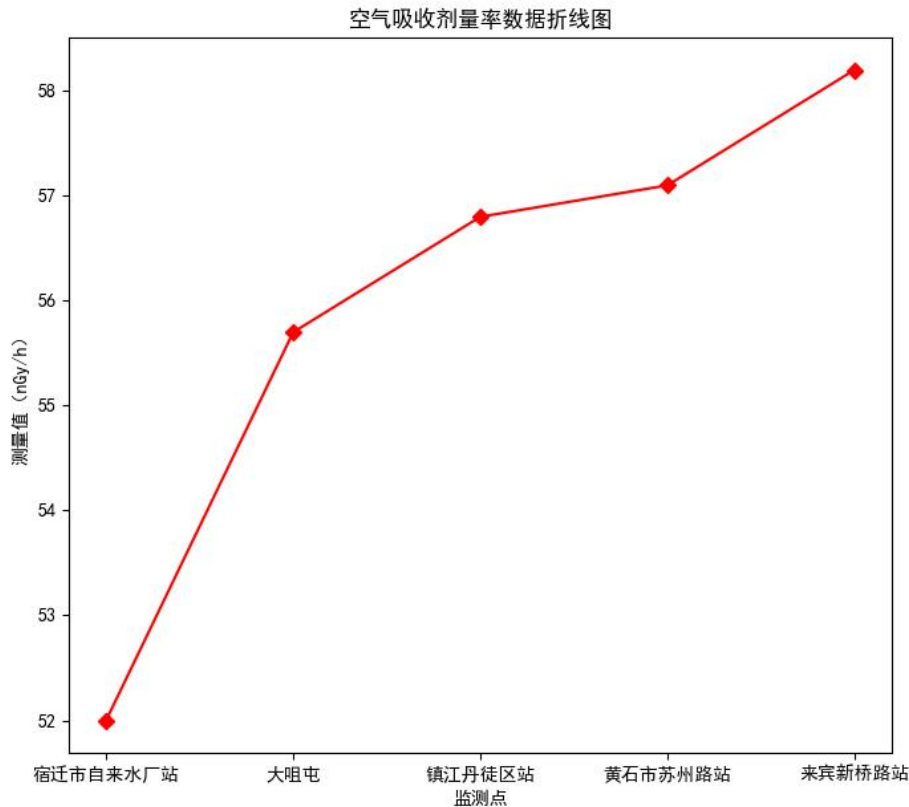


图 H2-4-2 空气吸收剂量率数据折线图

(2) 实施条件

抽测所需的软硬件如表 H2-4-1:

表 H2-4-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	Microsoft Office 2010 以上版本; MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019 安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pycharts1.x 等模块	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经	

项目	基本实施条件	备注
	验（副高及以上职称），或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书（2人/场）	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 80%，职业素养占该项目总分的 20%。具体评价标准见下表：

表 H2-4-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注	
工作任务 (80分)	数据采集	35分	新建、保存、运行项目和文件正确。	5分	1、考试舞弊、抄袭、没有按要求填写相关信息，本模块记0分。 2、严重违反考场纪律、造成恶劣影响的本项目记0分。
			认识 HTML 脚本语言，网页抓包分析正确，运行结果进行了截图。截图不正确扣 1 分。	10分	
			能根据实际项目需求进行问题分析，使用 Python 网络爬虫技术爬取网页中的数据。代码不正确 1 处扣 1 分。	10分	
			将数据写入 csv 文件。代码不正确 1 处扣 1 分。	10分	
	数据处 理和 析	25分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 2 分。	5分	
			read_file(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			print_info(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			hunan(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			sort(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
	数据可 视 化	20分	新建、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 1 分。	5分	
			导入 pandas、matplotlib 模块正确	2分	
			中文显示正确。	1分	
			读取 csv 文件正确。	2分	
设置标题正确。			2分		
按人次列升序排序正确。			2分		
绘制折线图正确。			2分		
X 轴标签和 Y 轴标签正确。			2分		
保存和显示图片正确。	2分				
职业 素养 (20分)	工 作 前 准 备	5分	做好工作前准备，检查电脑硬件（键盘、鼠标等），检查测试所需软件开发环境。不进行检查操作扣 5 分。	5分	
	专 业 素 养	10分	代码符合软件开发规范，命名规范，能做到见名知意；缩进统一，方便阅读；注释规范正确，至少要有 5 行注释，少 1 行或错 1 行扣 1 分。	10分	
	职 业 行 为 规 范	5分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5分	
总计		100			

评价内容	配分	评价要点	备注
	分		

项目 5：湖南城市 AQI 数据获取与应用

5.试题编号：2-5-1 湖南城市 AQI 数据获取与应用

(1) 任务描述

任务 1：数据采集

网页“<http://pm25.in/changsha>”中包含长沙 10 个监测点的空气质量监测数据，请编写程序抓取网页（网页样本保存在源素材文件夹下 src1 目录中）上的监测点、AQI、空气质量指数类别、PM2.5、PM10、CO，如表 H2-5-1 所示。将抓取的数据保存，以 bj20200721.csv 文件命名。

表 H2-5-1 长沙空气质量监测数据

监测点	AQI	空气质量指数类别
经开区环保局	28	优
.....

1) 使用 Pycharm 在考生文件夹下创建项目 task030501，在项目下创建 python 文件，命名为 task1.py。将源素材文件夹下的 src1 目录复制到项目 task030501 下。

2) 根据实际项目需求进行问题分析，书写注释。

3) 编写程序，正确爬取网页中的数据，运行程序并显示“已爬取”。

4) 编写程序，将抓取的数据保存在项目 task030501 下，以 cs20200728.csv 文件命名，运行程序并显示“已保存”。

5) 将运行结果截图，保存在考生文件夹下，以 task1.jpg 文件命名。

任务 2：数据处理和分析

使用 Python 的 pandas 读取 cs20200728.csv 文件（cs20200728.csv 文件保存在源素材文件夹下 task030502\src2 目录中），完成数据处理和分析。

1) 将源素材文件夹下 task030502 项目复制到考生文件夹下，使用 Pycharm 打开项目 task030502，打开 task2.py。

2) 根据实际项目需求进行问题分析，书写注释。

3) 编写程序，完成 read_file(filename)函数的定义，使用 pandas 读取 cs20200728.csv 文件中的数据，返回值类型为 DataFrame。

4) 编写程序，完成 print_info(filename)函数定义，显示 cs20200728.csv 文件中的内容，如图 H2-5-1 所示。

	监测点	AQI	空气质量指数类别	PM2.5	PM10	CO
0	经开区环保局	28	优	13	28	0.9
1	高开区环保局	69	良	50	60	1.0
2	马坡岭	44	优	16	44	0.8
3	湖南师范大学	39	优	22	39	0.9
4	雨花区环保局	29	优	15	29	0.7
5	伍家岭	32	优	14	32	0.5
6	火车新站	43	优	18	43	0.7
7	天心区环保局	28	优	16	28	0.5
8	湖南中医药大学	43	优	21	43	1.0
9	沙坪	_	NaN	_	29	0.8

图 H2-5-1 显示 cs20200728.csv 文件中的内容

5) 编写程序, 完成 `insert_date(filename)` 函数的定义, 将 2020-07-28 日期数据作为新的一列添加到 Excel 文件中, 列名为“监测日期”, 保存在项目 task030502 下, 文件命名为 cs20200728.xlsx。

6) 编写程序, 完成 `aqi(filename)` 函数的定义, 显示空气质量指数类别为优的记录。

7) 编写程序, 导入模块, 调用自定义的 `read_file`、`print_info`、`insert_date`、`aqi` 函数, 实现上述功能, 注意参数的正确设置。

8) 将运行结果截图, 保存在考生文件夹下, 以 task2.jpg 文件命名。

任务 3: 数据可视化

使用 Python 读取 hn20200728.csv 文件 (hn20200728.csv 文件保存在源素材文件夹下 src3 目录中) 中的数据, 完成数据处理和分析。

1) 使用 Pycharm 在考生文件夹下创建项目 task030503, 在项目下创建 python 文件, 命名为 task3.py。将源素材文件夹下的 src3 目录复制到项目 task030503 下。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 导入 pandas。使用 pandas 读取 hn20200728.csv 文件中的监测点和 AQI 数据。使用 matplotlib 绘制直方图, 如图 H2-5-2 所示。生成的文件保存在项目 task030503 下, 文件命名为 aqi.png。

4) 将运行结果截图, 保存在考生文件夹下, 以 task3.jpg 文件命名。

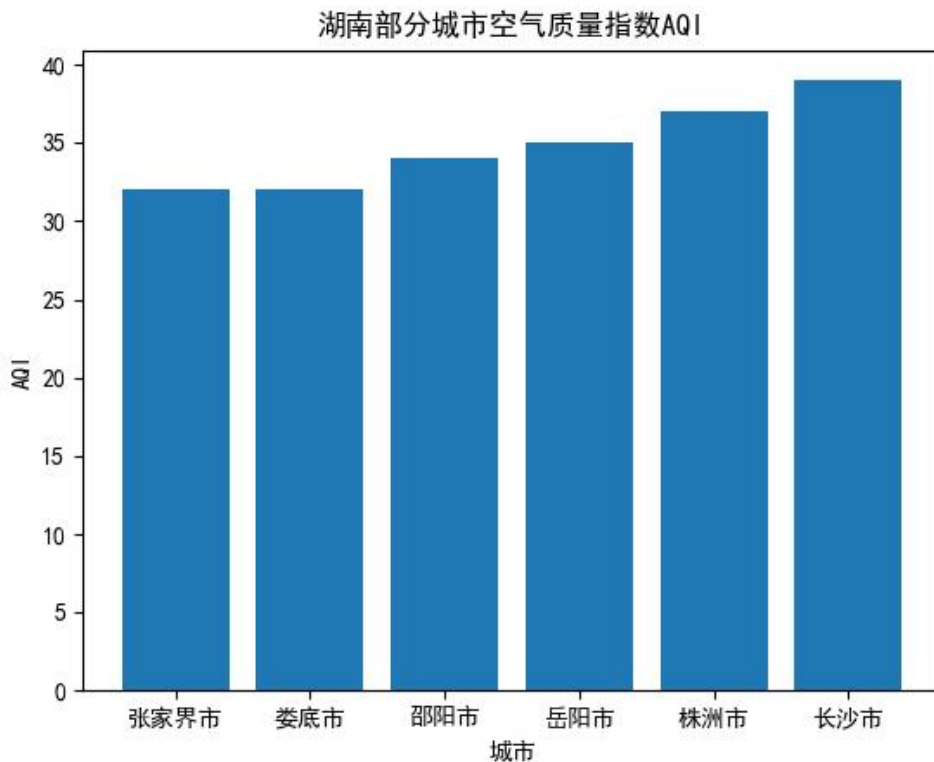


图 H2-5-2 湖南部分城市空气质量指数 AQI

(2) 实施条件

抽测所需的软硬件如表 H2-5-1:

表 H2-5-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	Microsoft Office 2010 以上版本; MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019 安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pyecharts1.x 等模块	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制, 评价内容包括工作任务、职业素养完成情况两个方面。其中, 工作任务完成质量占该项目总分的 80%, 职业素养占该项目总分的 20%。具体评价标准见下表:

表 H2-5-2 数据应用开发模块考核评分细则

评价内容	配分	评价要点	备注		
工作任务 (80 分)	数据 采集	新建、保存、运行项目和文件正确。	5 分	1、考试舞弊、抄袭、没有按要求填写相关信息, 本模块记 0 分。 2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。	
		认识 HTML 脚本语言, 网页抓包分析正确, 运行结果进行了截图。截图不正确扣 1 分。	10 分		
		能根据实际项目需求进行问题分析, 使用 Python 网络爬虫技术爬取网页中的数据。代码不正确 1 处扣 1 分。	10 分		
		将数据写入 csv 文件。代码不正确 1 处扣 1 分。	10 分		
	数据 处理和 析	25 分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图, 截图不正确扣 2 分。		5 分
			read_file(filename)函数定义正确。代码不正确 1 处扣 1 分。		5 分
			print_info(filename)函数定义正确。代码不正确 1 处扣 1 分。		5 分
			insert_date(filename)函数定义正确。代码不正确 1 处扣 1 分。		5 分
			aqi(filename)函数定义正确。代码不正确 1 处扣 1 分。		5 分
	数据 可视	20 分	新建、保存、运行项目和文件正确。运行结果进行了截图, 截图不正确扣 1 分。		5 分

评价内容		配分	评价要点	备注
化			导入 pandas、matplotlib 模块正确	2 分
			中文显示正确。	1 分
			读取 csv 文件正确。	2 分
			设置标题正确。	2 分
			按 AQI 列升序排序正确。	2 分
			绘制直方图正确。	2 分
			X 轴标签和 Y 轴标签正确。	2 分
			保存和显示图片正确。	2 分
职业素养 (20 分)	工作前准备	5 分	做好工作前准备, 检查电脑硬件(键盘、鼠标等), 检查测试所需软件开发环境。不进行检查操作扣 5 分。	5 分
	专业素养	10 分	代码符合软件开发规范, 命名规范, 能做到见名知意; 缩进统一, 方便阅读; 注释规范正确, 至少要有 5 行注释, 少 1 行或错 1 行扣 1 分。	10 分
	职业行为规范	5 分	着装干净、整洁。举止文明, 遵守考场纪律, 按顺序进出考场。考试过程中及结束后, 考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5 分
	总计	100 分		

项目 6: 地震历史数据获取与分析

6.试题编号: 4-6-1 地震历史数据获取与分析

(1) 任务描述

任务 1: 数据采集

网页“<http://ditu.92cha.com/dizhen.php>”中包含全球地震历史数据, 请编写程序抓取网页(网页样本保存在源素材文件夹下 src1 目录中)上的发震时刻、震级、经度、纬度、深度(千米)、参考位置, 如表 H2-6-1 所示。将抓取的数据保存, 以 earthquake.csv 文件命名。

表 H2-6-1 地震历史数据

发震时刻	震级	经度	纬度
2020-07-2822:38:17	3.8	86.8400	33.2700
.....

1) 使用 Pycharm 在考生文件夹下创建项目 task030601, 在项目下创建 python 文件, 命名为 task1.py。将源素材文件夹下的 src1 目录复制到项目 task030601 下。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 正确爬取网页中的数据, 运行程序并显示“已爬取”。

4) 编写程序, 将抓取的数据保存在项目 task030601 下, 以 earthquake.csv 文件命名, 运行程序并显示“已保存”。

5) 将运行结果截图, 保存在考生文件夹下, 以 task1.jpg 文件命名。

任务 2: 数据处理和分析

使用 Python 的 pandas 读取 earthquake.csv 文件 (earthquake.csv 文件保存在源素材文件夹下 task030602\src2 目录中), 完成数据处理和分析。

1) 将源素材文件夹下 task030602 项目复制到考生文件夹下, 使用 Pycharm 打开项目 task030602, 打开 task2.py。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 完成 `read_file(filename)` 函数的定义, 使用 `pandas` 读取 `earthquake.csv` 文件中的数据, 返回值类型为 `DataFrame`。

4) 编写程序, 完成 `print_info(filename)` 函数定义, 显示 `earthquake.csv` 文件中前 10 行的内容, 如图 H2-6-1 所示。

	发震时刻	震级	经度	纬度	深度(千米)	参考位置
0	2020-07-28 22:38:17	3.8	86.84	33.27	6.0	西藏那曲市尼玛县
1	2020-07-28 16:03:45	5.8	-161.30	54.90	40.0	美国阿拉斯加州以南海域
2	2020-07-28 11:32:47	3.9	94.89	30.35	10.0	西藏林芝市波密县
3	2020-07-28 09:46:40	2.4	119.11	31.54	7.0	江苏南京市溧水区(有感)
4	2020-07-28 03:18:13	3.1	95.04	31.91	8.0	西藏昌都市丁青县
5	2020-07-28 01:32:44	5.7	126.40	9.05	50.0	菲律宾棉兰老岛附近海域
6	2020-07-27 16:06:59	3.4	104.95	28.39	8.0	四川宜宾市长宁县
7	2020-07-27 13:50:18	3.4	94.83	30.43	10.0	西藏林芝市波密县
8	2020-07-27 13:14:48	5.3	104.70	20.90	10.0	越南
9	2020-07-26 20:52:27	5.5	122.48	24.27	50.0	台湾花莲县海域

图 H2-6-1 显示 `earthquake.csv` 文件前 10 行内容

5) 编写程序, 完成 `sort(filename)` 函数的定义, 按震级降序排序, 保存在项目 `task030602` 下, 文件命名为 `earthquake.xlsx`。

6) 编写程序, 完成 `select(filename)` 函数的定义, 显示震级在 5.0 以上的数据记录。

7) 编写程序, 导入模块, 调用自定义的 `read_file`、`print_info`、`sort`、`select` 函数, 实现上述功能, 注意参数的正确设置。

8) 将运行结果截图, 保存在考生文件夹下, 以 `task2.jpg` 文件命名。

任务 3: 数据可视化

使用 Python 读取 `earthquake.csv` 文件(`earthquake.csv` 文件保存在源素材文件夹下 `src3` 目录) 中的数据, 使用 `matplotlib` 绘制折线图, 显示震级最低的 5 条记录。

1) 使用 `Pycharm` 在考生文件夹下创建项目 `task030603`, 在项目下创建 `python` 文件, 命名为 `task3.py`。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 使用 `pandas` 读取 `earthquake.csv` 文件中的数据。使用 `matplotlib` 绘制折线图, 显示震级最低的 5 条记录, 如图 H2-6-2 所示。生成的文件保存在项目 `task030603` 下, 文件命名为 `quake.png`。

4) 运行结果截图, 保存在考生文件夹下, 以 `task3.jpg` 文件命名。

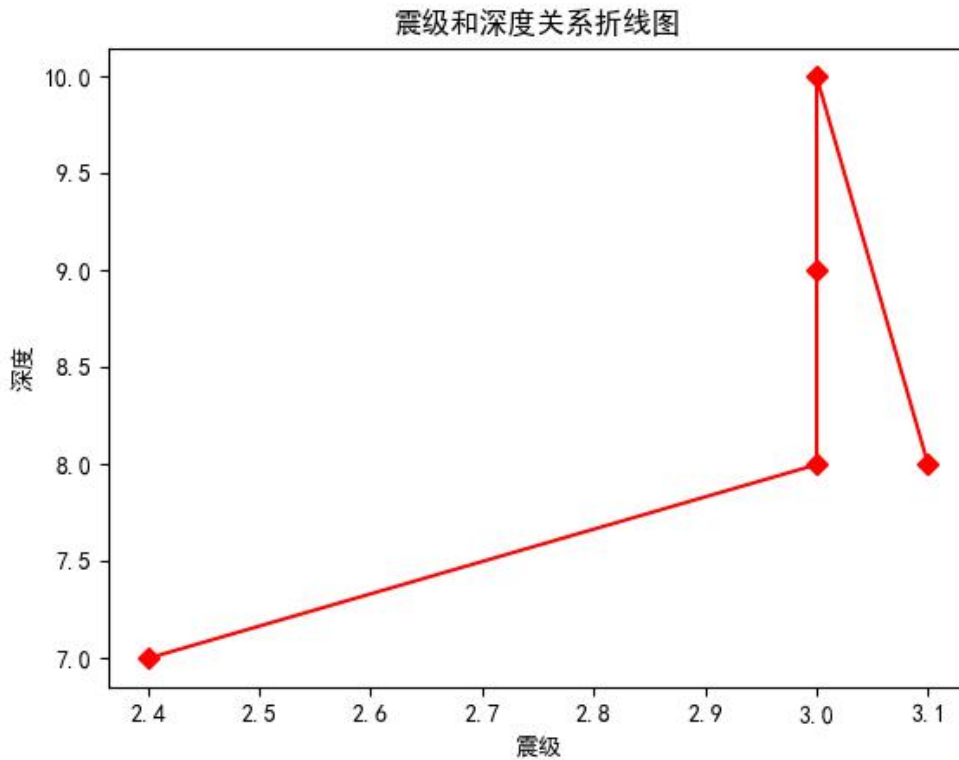


图 H2-6-2 震级和深度折线图

(2) 实施条件

抽测所需的软硬件如表 H2-6-1:

表 H2-6-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	Microsoft Office 2010 以上版本; MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019 安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pycharts1.x 等模块	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 80%，职业素养占该项目总分的 20%。具体评价标准见下表：

表 H2-6-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点		备注
工作任务 (80分)	数据采集	35分	新建、保存、运行项目和文件正确。	5分	1、考试舞弊、抄袭、没有按要求填写相关信息，本模块记0分。 2、严重违反考场纪律、造成恶劣影响的本项目记0分。
			认识 HTML 脚本语言，网页抓包分析正确，运行结果进行了截图。截图不正确扣 1 分。	10分	
			能根据实际项目需求进行问题分析，使用 Python 网络爬虫技术爬取网页中的数据。代码不正确 1 处扣 1 分。	10分	
			将数据写入 csv 文件。代码不正确 1 处扣 1 分。	10分	
	数据处理和析	25分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 2 分。	5分	
			read_file(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			print_info(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			sort(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			select(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
	数据可视化	20分	新建、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 1 分。	5分	
			导入 pandas、matplotlib 模块正确	2分	
			中文显示正确。	1分	
			读取 csv 文件正确。	2分	
			设置标题正确。	2分	
			按震级列升序排序正确。	2分	
绘制折线图正确。			2分		
X 轴标签和 Y 轴标签正确。			2分		
保存和显示图片正确。	2分				
职业素养 (20分)	工作前准备	5分	做好工作前准备，检查电脑硬件(键盘、鼠标等)，检查测试所需软件开发环境。不进行检查操作扣 5 分。	5分	
	专业素养	10分	代码符合软件开发规范，命名规范，能做到见名知意；缩进统一，方便阅读；注释规范正确，至少要有 5 行注释，少 1 行或错 1 行扣 1 分。	10分	
	职业行为规范	5分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5分	
	总计	100分			

项目 7：上海空气质量监测数据获取与分析

7.试题编号：4-7-1 上海空气质量监测数据获取与分析

(1) 任务描述

任务 1：数据采集

网页“<http://pm25.in/shanghai>”中包含上海 10 个监测点的空气质量监测数据，请编写程序抓取网页（网页样本保存在源素材文件夹下 src1 目录中）上的监测点、AQI、空气质量指数类别，如表 H2-7-1 所示。将抓取的数据保存，以 shanghai20200729.csv 文件命名。

表 H2-7-1 上海空气质量监测数据

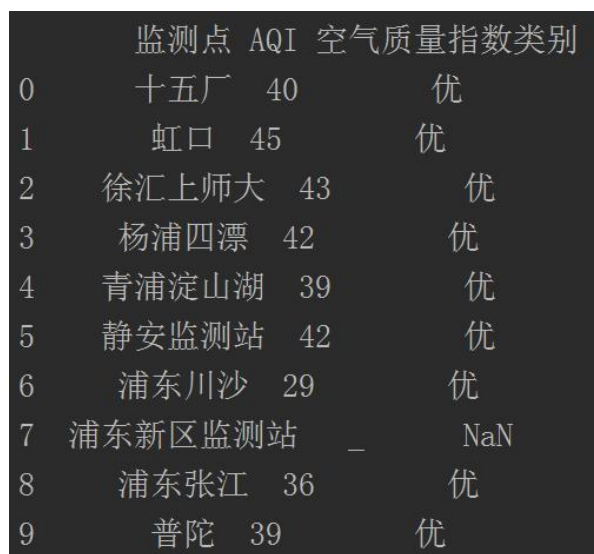
监测点	AQI	空气质量指数类别
十五厂	40	优
.....

- 1) 使用 Pycharm 在考生文件夹下创建项目 task030701，在项目下创建 python 文件，命名为 task1.py。将源素材文件夹下的 src1 目录复制到项目 task030701 下。
- 2) 根据实际项目需求进行问题分析，书写注释。
- 3) 编写程序，正确爬取网页中的数据，运行程序并显示“已爬取”。
- 4) 编写程序，将抓取的数据保存在项目 task030701 下，以 shanghai20200729.csv 文件命名，运行程序并显示“已保存”。
- 5) 将运行结果截图，保存在考生文件夹下，以 task1.jpg 文件命名。

任务 2：数据处理和分析

使用 Python 的 pandas 读取 shanghai20200729.csv 文件（shanghai20200729.csv 文件保存在源素材文件夹下 task030702\src2 目录中），完成数据处理和分析。

- 1) 将源素材文件夹下 task030702 项目复制到考生文件夹下，使用 Pycharm 打开项目 task030702，打开 task2.py。
- 2) 根据实际项目需求进行问题分析，书写注释。
- 3) 编写程序，完成 read_file(filename)函数的定义，使用 pandas 读取 shanghai20200729.csv 文件中的数据，返回值类型为 DataFrame。
- 4) 编写程序，完成 print_info(filename)函数定义，显示 shanghai20200729.csv 文件中的内容，如图 H2-7-1 所示。



	监测点	AQI	空气质量指数类别
0	十五厂	40	优
1	虹口	45	优
2	徐汇上师大	43	优
3	杨浦四漂	42	优
4	青浦淀山湖	39	优
5	静安监测站	42	优
6	浦东川沙	29	优
7	浦东新区监测站	_	NaN
8	浦东张江	36	优
9	普陀	39	优

图 H2-7-1 显示 shanghai20200729.csv.csv 文件中的内容

- 5) 编写程序，完成 insert_date(filename)函数的定义，将 2020-07-29 日期数

据作为新的一列添加到 Excel 文件中，保存在项目 task030702 下，文件命名为 shanghai20200729.xlsx。

6) 编写程序，完成 aqi(filename)函数的定义，显示空气质量指数类别为优的记录。

7) 编写程序，导入模块，调用自定义的 read_file、print_info、insert_date、aqi 函数，实现上述功能，注意参数的正确设置。

8) 将运行结果截图，保存在考生文件夹下，以 task2.jpg 文件命名。

任务 3：数据可视化

使用 Python 读取 shanghai20200729.csv 文件（shanghai20200729.csv 文件保存在源素材文件夹下 src3 目录中）中的数据，使用 matplotlib 绘制直方图，显示 aqi 最大的 5 个监测点。

1) 使用 Pycharm 在考生文件夹下创建项目 task030703，在项目下创建 python 文件，命名为 task3.py。将源素材文件夹下的 src3 目录复制到项目 task030703 下。

2) 根据实际项目需求进行问题分析，书写注释。

3) 编写程序，使用 pandas 读取 shanghai20200729.csv 文件中的数据。使用 matplotlib 绘制直方图，显示 AQI 最大的 5 个监测点，如图 H2-7-2 所示。生成的文件保存在项目 task030703 下，文件命名为 aqi.png。

4) 将运行结果截图，保存在考生文件夹下，以 task3.jpg 文件命名。

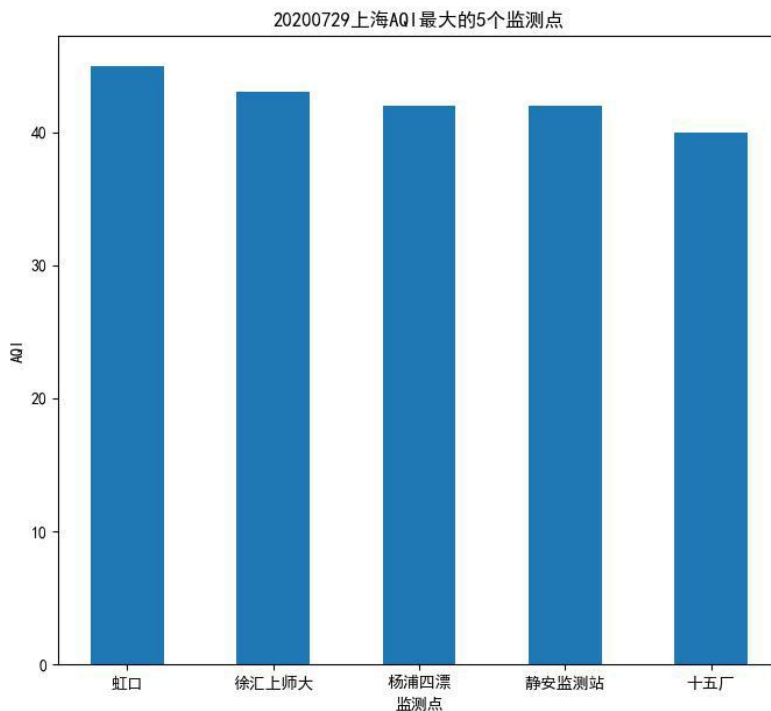


图 H2-7-2 AQI 最大的 5 个监测点直方图

(2) 实施条件

抽测所需的软硬件如表 H2-7-1:

表 H2-7-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB	用于保存测试人

项目	基本实施条件	备注
	以上, 硬盘空间 1T 以上);	员考试结果
工具	Microsoft Office 2010 以上版本; MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019 安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pycharts1.x 等模块	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验(工程师及以上职称)或从事本专业或相关专业具有 5 年以上的教学经验(副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书(2 人/场)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验(工程师及以上职称)或从事本专业或相关专业具有 5 年以上的教学经验(副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书(2 人/场)	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制, 评价内容包括工作任务、职业素养完成情况两个方面。其中, 工作任务完成质量占该项目总分的 80%, 职业素养占该项目总分的 20%。具体评价标准见下表:

表 H2-7-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注	
工作任务 (80分)	数据采集	35分	新建、保存、运行项目和文件正确。	5分	1、考试舞弊、抄袭、没有按 要求填写相关 信息, 本模块 记 0 分。 2、严重违反 考场纪律、造 成恶劣影响的 本项记 0 分。
			认识 HTML 脚本语言, 网页抓包分析正确, 运行结果进行了截图。截图不正确扣 1 分。	10分	
			能根据实际项目需求进行问题分析, 使用 Python 网络爬虫技术爬取网页中的数据。代码不正确 1 处扣 1 分。	10分	
			将数据写入 csv 文件。代码不正确 1 处扣 1 分。	10分	
	数据处理 和析	25分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图, 截图不正确扣 2 分。	5分	
			read_file(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			print_info(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			insert_date(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			aqi(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
	数据 可视 化	20分	新建、保存、运行项目和文件正确。运行结果进行了截图, 截图不正确扣 1 分。	5分	
			导入 pandas、matplotlib 模块正确	2分	
			中文显示正确。	1分	
			读取 csv 文件正确。	2分	
设置标题正确。			2分		

评价内容		配分	评价要点	备注
			按 AQI 列降序排序正确。	2 分
			绘制直方图正确。	2 分
			X 轴标签和 Y 轴标签正确。	2 分
			保存和显示图片正确。	2 分
职业素养 (20 分)	工作前准备	5 分	做好工作前准备, 检查电脑硬件(键盘、鼠标等), 检查测试所需软件开发环境。不进行检查操作扣 5 分。	5 分
	专业素养	10 分	代码符合软件开发规范, 命名规范, 能做到见名知意; 缩进统一, 方便阅读; 注释规范正确, 至少要有 5 行注释, 少 1 行或错 1 行扣 1 分。	10 分
	职业行为规范	5 分	着装干净、整洁。举止文明, 遵守考场纪律, 按顺序进出考场。考试过程中及结束后, 考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5 分
	总计	100 分		

项目 8: 杭州空气质量监测数据获取与分析

8.试题编号: 4-8-1 杭州空气质量监测数据获取与分析

(1) 任务描述

任务 1: 数据采集

网页“<http://pm25.in/hangzhou>”中包含杭州 11 个监测点的空气质量监测数据, 请编写程序抓取网页(网页样本保存在源素材文件夹下 src1 目录中)上的监测点、AQI、空气质量指数类别, 如表 H2-8-1 所示。将抓取的数据保存, 以 hz20200729.xlsx 文件命名。

表 H2-8-1 杭州空气质量监测数据

监测点	AQI	空气质量指数类别
滨江	68	良
.....

1) 使用 Pycharm 在考生文件夹下创建项目 task030801, 在项目下创建 python 文件, 命名为 task1.py。将源素材文件夹下的 src1 目录复制到项目 task030801 下。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 正确爬取网页中的数据, 运行程序并显示“已爬取”。

4) 编写程序, 将抓取的数据保存在项目 task030801 下, 以 hz20200729.xlsx 文件命名, 运行程序并显示“已保存”。

5) 将运行结果截图, 保存在考生文件夹下, 以 task1.jpg 文件命名。

任务 2: 数据处理和分析

使用 Python 的 pandas 读取 hz20200729.csv 文件(hz20200729.csv 文件保存在源素材文件夹下 task030802\src2 目录中), 完成数据处理和分析。

1) 将源素材文件夹下 task030802 项目复制到考生文件夹下, 使用 Pycharm 打开项目 task030802, 打开 task2.py。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 完成 read_file(filename)函数的定义, 使用 pandas 读取 hz20200729.csv 文件中的数据, 返回值类型为 DataFrame。

4) 编写程序, 完成 print_info(filename)函数定义, 显示 hz20200729.csv 文件中的内容, 如图 H2-8-1 所示。

	监测点	AQI	空气质量指数类别
0	滨江	68	良
1	西溪	62	良
2	千岛湖	42	优
3	下沙	48	优
4	卧龙桥	59	良
5	浙江农大	53	良
6	朝晖五区	87	良
7	和睦小学	97	良
8	临平镇	68	良
9	城厢镇	47	优
10	云栖	-	NaN

图 H2-8-1 显示 bj20200721.csv 文件中的内容

5) 编写程序, 完成 `insert_date(filename)` 函数的定义, 将 2020-07-29 日期数据作为新的一列添加到文件中, 保存在项目 task030802 下, 文件命名为 hz20200729.xlsx。

6) 编写程序, 完成 `aqi(filename)` 函数的定义, 显示空气质量指数类别为良的记录。

7) 编写程序, 导入模块, 调用自定义的 `read_file`、`print_info`、`insert_date`、`aqi` 函数, 实现上述功能, 注意参数的正确设置。

8) 将运行结果截图, 保存在考生文件夹下, 以 task2.jpg 文件命名。

任务 3: 数据可视化

使用 Python 读取 hz20200729.csv 文件 (hz20200729.csv 文件保存在源素材文件夹下 src3 目录中) 中的数据, 使用 matplotlib 绘制折线图, 显示各监测点 AQI 值。

1) 使用 Pycharm 在考生文件夹下创建项目 task030803, 在项目下创建 python 文件, 命名为 task3.py。将源素材文件夹下的 src3 目录复制到项目 task030803 下。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 使用 pandas 读取 hz20200729.csv 文件中的数据。

4) 使用 matplotlib 绘制折线图, 显示各监测点 AQI 值, 如图 H2-8-2 所示。生成的文件保存在项目 task030803 下, 文件命名为 aqi.png。

5) 将运行结果截图, 保存在考生文件夹下, 以 task3.jpg 文件命名。

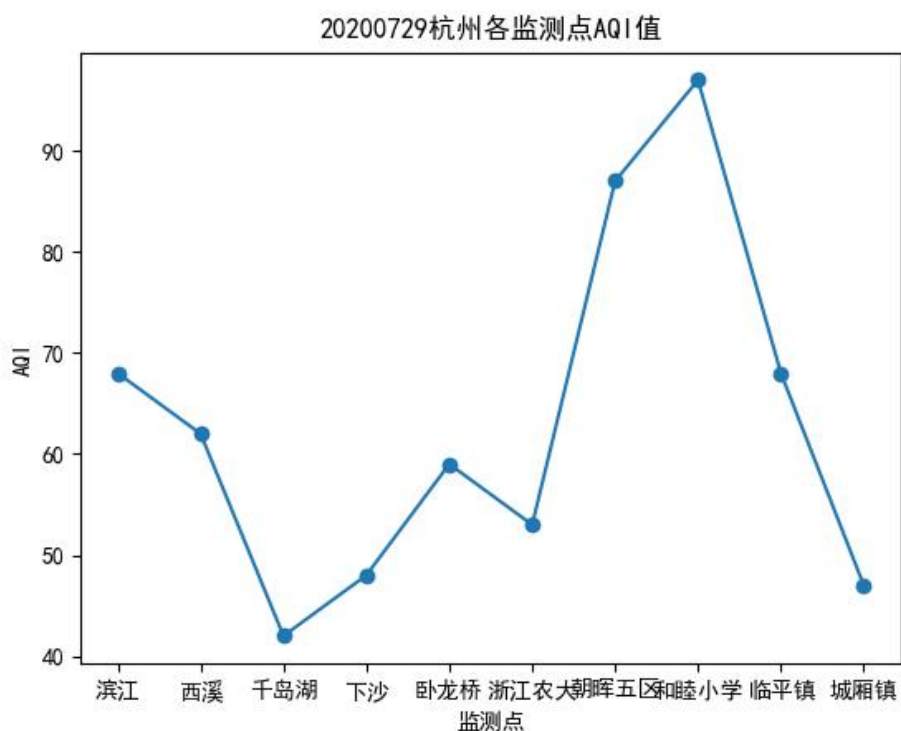


图 H2-8-2 杭州各监测点 AQI 值折线图

(2) 实施条件

抽测所需的软硬件如表 H2-8-1:

表 H2-8-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	Microsoft Office 2010 以上版本; MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019 安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pycharts1.x 等模块	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 80%，职业素养占该项目总分的 20%。具体评价标准见下表：

表 H2-8-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注	
工作任务 (80分)	数据采集	35分	新建、保存、运行项目和文件正确。	5分	
			认识 HTML 脚本语言，网页抓包分析正确，运行结果进行了截图。截图不正确扣 1 分。	10分	
			能根据实际项目需求进行问题分析，使用 Python 网络爬虫技术爬取网页中的数据。代码不正确 1 处扣 1 分。	10分	
			将数据写入 csv 文件。代码不正确 1 处扣 1 分。	10分	
	数据处理和析	25分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 2 分。	5分	
			read_file(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			print_info(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			insert_date(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			aqi(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
	数据可视化	20分	新建、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 1 分。	5分	
			导入 pandas、matplotlib 模块正确	2分	
			中文显示正确。	1分	
			读取 csv 文件正确。	2分	
			设置标题正确。	3分	
			绘制折线图正确。	3分	
			X 轴标签和 Y 轴标签正确。	2分	
	保存和显示图片正确。	2分			
	职业素养 (20分)	工作前准备	5分	做好工作前准备，检查电脑硬件(键盘、鼠标等)，检查测试所需软件开发环境。不进行检查操作扣 5 分。	5分
		专业素养	10分	代码符合软件开发规范，命名规范，能做到见名知意；缩进统一，方便阅读；注释规范正确，至少要有 5 行注释，少 1 行或错 1 行扣 1 分。	10分
职业行为规范		5分	着装干净、整洁。举止文明，遵守考场纪律，按顺序进出考场。考试过程中及结束后，考试桌面及地面不符合 6S 基本要求的扣 3—5 分。	5分	
	总计	100分			

1、考试舞弊、抄袭、没有按要求填写相关信息，本模块记 0 分。
2、严重违反考场纪律、造成恶劣影响的本项目记 0 分。

项目 9：辐射监测数据获取与分析

9.试题编号：4-9-1 辐射监测数据获取与分析

(1) 任务描述

任务 1：数据采集

网页“<http://ditu.92cha.com/kqxsjll.php>”中包含辐射监测中空气吸收剂量率实时数据，请编写程序抓取网页（网页样本保存在源素材文件夹下 src1 目录中）上的省份、监测点、测量时间、测量值（nGy/h），如表 H2-9-1 所示。将抓取的数据保存，以 fs20200729.csv 文件命名。

表 H2-9-1 空气吸收剂量率数据

省份	监测点	测量时间	测量值（nGy/h）
北京	奥林匹克森林公园站	2020-07-2911:04	76.7
.....

1) 使用 Pycharm 在考生文件夹下创建项目 task030901，在项目下创建 python 文件，命名为 task1.py。将源素材文件夹下的 src1 目录复制到项目 task030901 下。

2) 根据实际项目需求进行问题分析，书写注释。

3) 编写程序，正确爬取网页中的数据，运行程序并显示“已爬取”。

4) 编写程序，将抓取的数据保存在项目 task030901 下，以 fs20200729.csv 文件命名，运行程序并显示“已保存”。

5) 将运行结果截图，保存在考生文件夹下，以 task1.jpg 文件命名。

任务 2：数据处理和分析

使用 Python 的 pandas 读取 fs20200729.csv 文件（fs20200729.csv 文件保存在源素材文件夹下 task030902\src2 目录中），完成数据处理和分析。

1) 将源素材文件夹下 task030902 项目复制到考生文件夹下，使用 Pycharm 打开项目 task030902，打开 task2.py。

2) 根据实际项目需求进行问题分析，书写注释。

3) 编写程序，完成 read_file(filename) 函数的定义，使用 pandas 读取 fs20200729.csv 文件中的数据，返回值类型为 DataFrame。

4) 编写程序，完成 print_info(filename) 函数定义，显示 fs20200729.csv 文件中前 8 行的内容，如图 H2-9-1 所示。

	省份	监测点	测量时间	测量值（nGy/h）
0	北京	奥林匹克森林公园站	2020-07-29 11:04	76.7
1	北京	北京万柳中路站	2020-07-29 11:04	90.0
2	北京	北京门头沟西苑路站	2020-07-29 11:04	63.9
3	北京	北京昌平昌赤路站	2020-07-29 11:04	68.4
4	北京	亦庄博大公园站	2020-07-29 11:04	72.5
5	北京	北京平谷赵赵路站	2020-07-29 11:04	66.4
6	北京	北京延庆夏都公园站	2020-07-29 11:04	67.7
7	天津	南开复康路站	2020-07-29 11:04	72.6

图 H2-9-1 显示 fs20200729.csv 文件前 8 行内容

5) 编写程序，完成 gd(filename) 函数的定义，将广东监测点的数据保存在项目 task030902 下，文件命名为 gd.xlsx。

6) 编写程序，完成 sort(filename) 函数的定义，按测量值（nGy/h）从大到小

排序，显示运行结果。

7) 编写程序，导入模块，调用自定义的 `read_file`、`print_info`、`gd`、`sort` 函数，实现上述功能，注意参数的正确设置。

8) 将运行结果截图，保存在考生文件夹下，以 `task2.jpg` 文件命名。

任务 3：数据可视化

使用 Python 读取 `gd.csv` 文件（`gd.csv` 文件保存在源素材文件夹下 `src3` 目录中）中的数据，使用 `matplotlib` 绘制折线图。

1) 使用 Pycharm 在考生文件夹下创建项目 `task030903`，在项目下创建 `python` 文件，命名为 `task3.py`。将源素材文件夹下的 `src3` 目录复制到项目 `task030903` 下。

2) 根据实际项目需求进行问题分析，书写注释。

3) 编写程序，使用 `pandas` 读取 `gd.csv` 文件中的数据。

4) 使用 `matplotlib` 绘制折线图，显示测量值（`nGy/h`）最低的 5 条数据分布情况，如图 H2-9-2 所示。生成的文件保存在项目 `task030903` 下，文件命名为 `gd.png`。

5) 将运行结果截图，保存在考生文件夹下，以 `task3.jpg` 文件命名。

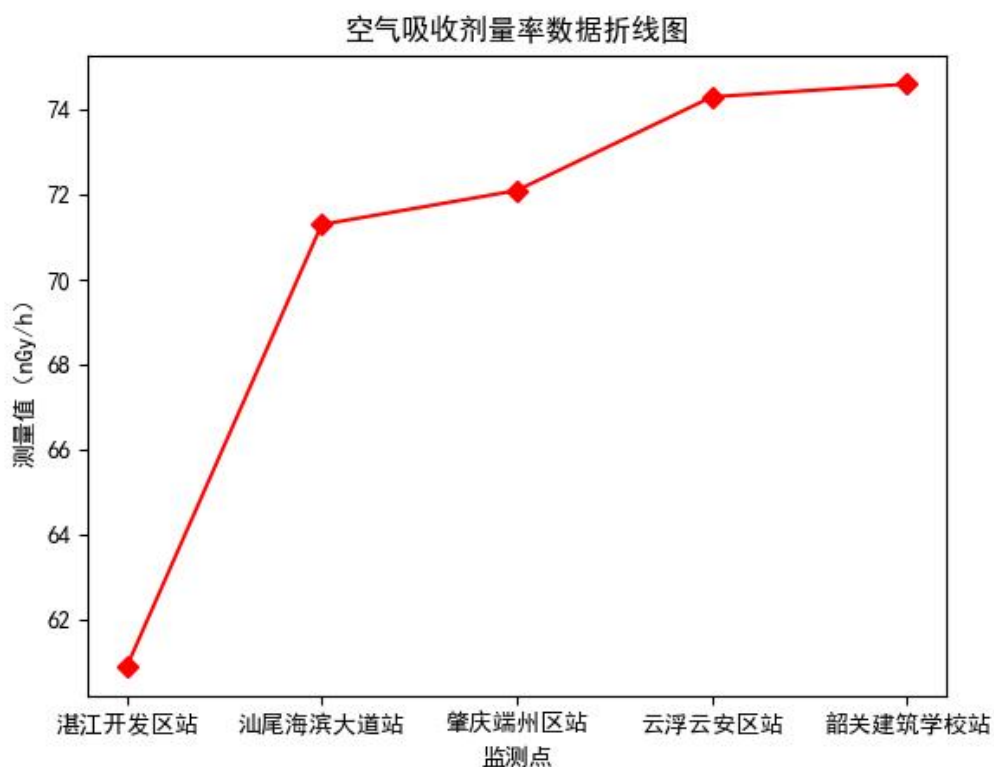


图 H2-9-2 空气吸收剂量率数据折线图

(2) 实施条件

抽测所需的软硬件如表 H2-9-1:

表 H2-9-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果

项目	基本实施条件	备注
工具	Microsoft Office 2010 以上版本； MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019 安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pycharts1.x 等模块	参考人员使用软件进行工作
测评专家	现场测评专家：在本行业具有 3 年以上从业经验（工程师及以上职称）或从事本专业或相关专业具有 5 年以上的教学经验（副高及以上职称），或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书（2 人/场）	测评专家满足任一条件
	结果测评专家：在本行业具有 3 年以上从业经验（工程师及以上职称）或从事本专业或相关专业具有 5 年以上的教学经验（副高及以上职称），或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书（2 人/场）	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制，评价内容包括工作任务、职业素养完成情况两个方面。其中，工作任务完成质量占该项目总分的 80%，职业素养占该项目总分的 20%。具体评价标准见下表：

表 H2-9-2 数据应用开发模块考核评分细则

评价内容		配分	评价要点	备注	
工作任务 (80分)	数据采集	35分	新建、保存、运行项目和文件正确。	5分	1、考试舞弊、抄袭、没有按要求填写相关信息，本模块记0分。 2、严重违反考场纪律、造成恶劣影响的本项目记0分。
			认识 HTML 脚本语言，网页抓包分析正确，运行结果进行了截图。截图不正确扣 1 分。	10分	
			能根据实际项目需求进行问题分析，使用 Python 网络爬虫技术爬取网页中的数据。代码不正确 1 处扣 1 分。	10分	
			将数据写入 csv 文件。代码不正确 1 处扣 1 分。	10分	
	数据处理和析	25分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 2 分。	5分	
			read_file(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			print_info(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			gd(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
			sort(filename)函数定义正确。代码不正确 1 处扣 1 分。	5分	
	数据可视化	20分	新建、保存、运行项目和文件正确。运行结果进行了截图，截图不正确扣 1 分。	5分	
			导入 pandas、matplotlib 模块正确	2分	
			中文显示正确。	1分	
			读取 csv 文件正确。	2分	
			设置标题正确。	2分	
			按测量值列升序排序正确。	2分	

评价内容		配分	评价要点	备注
			绘制折线图正确。	2分
			X轴标签和Y轴标签正确。	2分
			保存和显示图片正确。	2分
职业素养 (20分)	工作前准备	5分	做好工作前准备,检查电脑硬件(键盘、鼠标等),检查测试所需软件开发环境。不进行检查操作扣5分。	5分
	专业素养	10分	代码符合软件开发规范,命名规范,能做到见名知意;缩进统一,方便阅读;注释规范正确,至少要有5行注释,少1行或错1行扣1分。	10分
	职业行为规范	5分	着装干净、整洁。举止文明,遵守考场纪律,按顺序进出考场。考试过程中及结束后,考试桌面及地面不符合6S基本要求的扣3—5分。	5分
	总计	100分		

项目 10: 广东城市 AQI 数据获取与应用

10.试题编号: 4-10-1 广东城市 AQI 数据获取与应用

(1) 任务描述

任务 1: 数据采集

网页“<http://pm25.in/guangzhou>”中包含广州 12 个监测点的空气质量监测数据,请编写程序抓取网页(网页样本保存在源素材文件夹下 src1 目录中)上的监测点、AQI、空气质量指数类别、首要污染物、PM2.5、PM10、CO,如表 H2-10-1 所示。将抓取的数据保存,以 gz20200729.csv 文件命名。

表 H2-10-1 广州空气质量监测数据

监测点	AQI	空气质量指数类别
广雅中学	27	优
.....

1) 使用 Pycharm 在考生文件夹下创建项目 task031001,在项目下创建 python 文件,命名为 task1.py。将源素材文件夹下的 src1 目录复制到项目 task031001 下。

2) 根据实际项目需求进行问题分析,书写注释。

3) 编写程序,正确爬取网页中的数据,运行程序并显示“已爬取”。

4) 编写程序,将抓取的数据保存在项目 task031001 下,以 gz20200729.csv 文件命名,运行程序并显示“已保存”。

5) 将运行结果截图,保存在考生文件夹下,以 task1.jpg 文件命名。

任务 2: 数据处理和分析

使用 Python 的 pandas 读取 gz20200729.csv 文件(gz20200729.csv 文件保存在源素材文件夹下 task031002\src2 目录中),完成数据处理和分析。

1) 将源素材文件夹下 task031002 项目复制到考生文件夹下,使用 Pycharm 打开项目 task031002,打开 task2.py。

2) 根据实际项目需求进行问题分析,书写注释。

3) 编写程序,完成 read_file(filename)函数的定义,使用 pandas 读取 gz20200729.csv 文件中的数据,返回值类型为 DataFrame。

4) 编写程序,完成 print_info(filename)函数定义,显示 gz20200729.csv 文件中的内容,如图 H2-10-1 所示。

	监测点	AQI	空气质量指数类别	首要污染物	PM2.5	PM10	CO
0	广雅中学	27	优	—	11	27	0.7
1	市五中	25	优	—	10	25	0.6
2	天河职幼	—	NaN	—	—	—	—
3	广东商学院	29	优	—	10	29	0.7
4	市八十六中	32	优	—	11	32	0.6
5	番禺中学	24	优	—	8	24	0.5
6	花都师范	30	优	—	10	30	0.5
7	市监测站	22	优	—	7	22	0.5
8	九龙镇镇龙	31	优	—	7	31	0.6
9	麓湖	20	优	—	6	20	0.7
10	帽峰山森林公园	23	优	—	9	23	0.5
11	体育西	28	优	—	11	28	0.6

图 H2-10-1 显示 gz20200729.csv 文件中的内容

5) 编写程序, 完成 `insert_date(filename)` 函数的定义, 将 2020-07-29 日期数据作为新的一列添加到 Excel 文件中, 保存在项目 task031002 下, 文件命名为 gz20200729.xlsx。

6) 编写程序, 完成 `wr(filename)` 函数的定义, 显示 PM10 大于 10 的记录。

7) 编写程序, 导入模块, 调用自定义的 `read_file`、`print_info`、`insert_date`、`wr` 函数, 实现上述功能, 注意参数的正确设置。

8) 将运行结果截图, 保存在考生文件夹下, 以 task2.jpg 文件命名。

任务 3: 数据可视化

使用 Python 读取 gd20200729.csv 文件 (gd20200729.csv 文件保存在源素材文件夹下 src3 目录中) 中的数据, 使用 matplotlib 绘制直方图。

1) 使用 Pycharm 在考生文件夹下创建项目 task031003, 在项目下创建 python 文件, 命名为 task3.py。将源素材文件夹下的 src3 目录复制到项目 task031003 下。

2) 根据实际项目需求进行问题分析, 书写注释。

3) 编写程序, 使用 pandas 读取 bj20200721.csv 文件中的数据。使用 matplotlib 绘制直方图, 显示 AQI 最少的 5 个监测点, 如图 H2-10-2 所示。生成的文件保存在项目 task031003 下, 文件命名为 aqi.png

5) 将运行结果截图, 保存在考生文件夹下, 以 task3.jpg 文件命名。

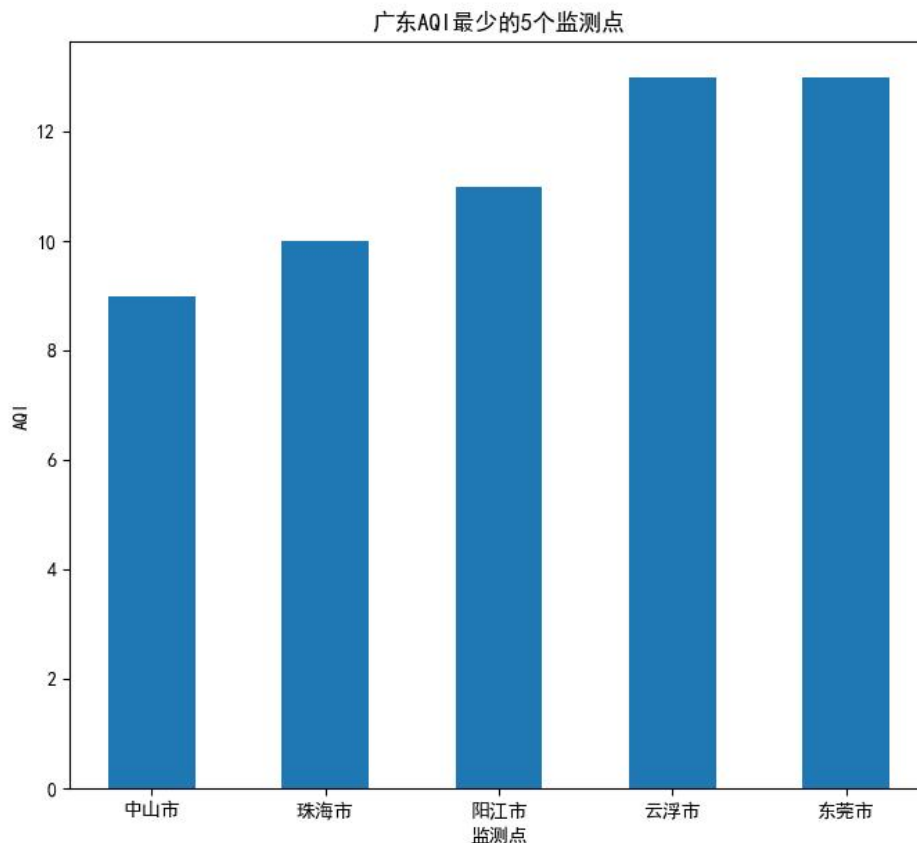


图 H2-10-2 广东 AQI 最少五个监测点

(2) 实施条件

抽测所需的软硬件如表 H2-10-1:

表 H2-10-1 数据应用开发模块实施条件

项目	基本实施条件	备注
场地	能同时容纳 50 人以上现场考核	
设备	50 台以上的主流计算机安装 Windows7 或更高版本考生计算机 (CPU2.0GHZ 以上, 内存 8GB, 硬盘空间 500G 以上);	每人一台
	FTP 服务器 1 台考场服务器 (CPU3.0GHZ 以上, 内存 32GB 以上, 硬盘空间 1T 以上);	用于保存测试人员考试结果
工具	Microsoft Office 2010 以上版本; MySQL 5.6+ Navicate for MySQL 7 及以上版本 Python3.6 Pycharm2019 安装 requests、beautifulsoup4、urllib3、numpy、pandas、matplotlib、pyecharts1.x 等模块	参考人员使用软件进行工作
测评专家	现场测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经验 (副高及以上职称), 或具有软件设计师、系统分析师、Web 前端开发工程师等资格证书 (2 人/场)	测评专家满足任一条件
	结果测评专家: 在本行业具有 3 年以上从业经验 (工程师及以上职称) 或从事本专业或相关专业具有 5 年以上的教学经	

项目	基本实施条件	备注
	验(副高及以上职称),或具有软件设计师、系统分析师、Web前端开发工程师等资格证书(2人/场)	

(3) 考核时量

考核时长 180 分钟。

(4) 评分细则

数据应用开发模块的考核实行 100 分制,评价内容包括工作任务、职业素养完成情况两个方面。其中,工作任务完成质量占该项目总分的 80%,职业素养占该项目总分的 20%。具体评价标准见下表:

表 H2-10-2 数据应用开发模块考核评分细则

评价内容	配分	评价要点	备注	
工作任务 (80分)	数据 采集	35分	新建、保存、运行项目和文件正确。5分	1、考试舞弊、抄袭、没有按要求填写相关信息,本模块记0分。 2、严重违反考场纪律、造成恶劣影响的本项目记0分。
			认识 HTML 脚本语言,网页抓包分析正确,运行结果进行了截图。截图不正确扣1分。10分	
			能根据实际项目需求进行问题分析,使用 Python 网络爬虫技术爬取网页中的数据。代码不正确1处扣1分。10分	
			将数据写入 csv 文件。代码不正确1处扣1分。10分	
	数据 处理和 分析	25分	复制、打开、保存、运行项目和文件正确。运行结果进行了截图,截图不正确扣2分。5分	
			read_file(filename)函数定义正确。代码不正确1处扣1分。5分	
			print_info(filename)函数定义正确。代码不正确1处扣1分。5分	
			insert_date(filename)函数定义正确。代码不正确1处扣1分。5分	
			wr(filename)函数定义正确。代码不正确1处扣1分。5分	
	数据 可视 化	20分	新建、保存、运行项目和文件正确。运行结果进行了截图,截图不正确扣1分。5分	
			导入 pandas、matplotlib 模块正确2分	
			中文显示正确。1分	
			读取 csv 文件正确。2分	
			设置标题正确。2分	
			按 AQI 列升序排序正确。2分	
			绘制直方图正确。2分	
			X 轴标签和 Y 轴标签正确。2分	
	保存和显示图片正确。2分			
	职业 素养 (20分)	工作 前 准备	5分	
专业 素养		10分	代码符合软件开发规范,命名规范,能做到见名知意;缩进统一,方便阅读;注释规范正确,至少要有5行注释,少1行或错1行扣1分。10分	
职业 行为 规范		5分	着装干净、整洁。举止文明,遵守考场纪律,按顺序进出考场。考试过程中及结束后,考试桌面及地面不符合6S基本要求的扣3—5分。5分	

评价内容		配分	评价要点	备注
	总计	100分		