

普通专升本免试入学专业综合知识答辩测试大纲

（数据科学与大数据技术专业）

一、测试目标

专业综合知识问答测试充分体现数据科学与大数据技术专业的人才培养目标与要求，全面考查学生对专业基础知识和基本技能的掌握情况，特别是对学生的思维、反应、表达、创新、综合素质、思想政治素质及专业知识综合运用能力和分析解决问题能力的考查。

二、命题的指导思想和原则

以理解、掌握和综合运用的能力作为命题的指导思想，每个考核知识点的题目着重体现综合能力。命题的原则按照测试考核知识点内容确定问答题目数，问答题目数量不少于 15 个，每问难度要保持一致，题量应与限定时间相匹配，问答题目表达要简练、明了、准确。

三、测试方式

专业综合知识答辩实行在公布的测试考核知识点范围内随机抽题，每人抽取一题，现场作答。

四、测试时间

每生测试时长为 15 分钟。

五、试卷结构

测试总分为 100 分，每个考核知识点的题目应能够体现语言表达（10 分）、应试态度（10 分）、知识内容（30 分）、专业水平及综合能力（30 分）、计算机应用能力（20 分）五部分内容。

六、参考书目

- 1.《大数据导论》，刘鹏主编，清华大学出版社，2018 年
- 2.《Linux 系统应用基础教程》，张小进主编，机械工业出版社，2020 年
- 3.《Python 程序设计基础》，董付国主编，清华大学出版社，2018 年

4.《Java 语言程序设计》，郑莉，张宇主编，清华大学出版社，2021年

七、测试的基本要求

1.本专业测试大纲适用于计算机专业大类下数据科学与大数据技术专业的普通专升本免试入学专业综合知识问答测试。

2.本门测试主要考核学生大数据技术应用、数据科学理论知识的专业综合能力。

3.考生需按照公布的测试考核知识点内容及数量自行准备测试应答资料。资料内容中，文案资料格式为 PDF 文档通用格式，演示内容为 PowerPoint 演示文稿的 PPT 文档格式。

4.考生需在测试当日将全部资料提交给抽签室的工作人员，由工作人员按照考生所抽题目将其对应的应试资料交主考室。

5.考生的应试资料中不得出现涵盖有标识考生姓名、考号、毕业学校等明显的标示性内容，如果发现按作弊处理。

八、测试考核知识点范围

考核知识点及要求：

考核知识点 1 简述 Linux 的几个运行级别及其相应的含义，并使用 Shell 命令(systemd)切换运行级别。

考核知识点 2 编写 Shell 脚本程序“compare.sh”，实现从终端输入任意两个数，比较两个数大小，并输出其比较结果。

考核知识点 3 使用 Shell 命令完成以下目录和文件的相关操作。

1) 创建一个名为 `students` 的用户组。

2) 创建一个名为 `student` 的用户，并将其添加到 `students` 用户组中。

3) 在用户的主目录下创建一个名为 `test` 的目录，并进入该目录。确保 `test` 目录具有读、写、执行权限。

4) 在 `test` 目录中创建两个文件，分别命名为 `file1.txt` 和

`file2.txt`。给 `file1.txt` 赋予读、写、执行权限，给 `file2.txt` 赋予只读权限。确保这两个文件的所有者是 `student`，用户组是 `students`。

5) 在 `test` 目录中创建一个名为 `sub` 的子目录，并进入该子目录。给 `sub` 目录赋予读、写、执行权限。确保这个子目录的所有者是 `student`，用户组是 `students`。

考核知识点 4 请编写一个 Python 函数，实现给定一个正整数 n ，返回 n 的阶乘。

考核知识点 5 编写一个 Python 程序，统计多个文本文件中的词频数，具体要求如下：

- 1) 函数接收一个文件路径列表作为参数。
- 2) 使用 `try-except` 结构处理文件打开异常。
- 3) 读入每个文件内容，利用正则表达式或字符串处理获得单词列表。
- 4) 使用字典统计每个单词的出现次数。
- 5) 返回词频字典。
- 6) 在 `main` 中调用函数。

考核知识点 6 编写一个 Python 程序，实现一个简单的学生信息管理系统。该系统具有以下基本功能：

1) 使用一个字典来存储学生信息，其中键是学生的学号，值是包含学生姓名和成绩的列表。

2) 提供以下操作选项：添加学生信息（学号、姓名、成绩）；查找学生信息，根据学号查询学生姓名和成绩；显示所有学生信息；计算并显示学生的平均成绩；退出系统。

3) 使用一个循环，使用户能够多次选择操作，直到选择退出为止。

考核知识点 7 有以下数字列表 `numbers = [5, 2, 6, 8, 3, 9]`，请完成以下操作：

- 1) 计算列表元素之和。
- 2) 找出列表中的最大值和最小值。

- 3) 将列表转换为元组。
- 4) 创建一个字典，键是列表中的元素，值是该元素的平方。
- 5) 打印输出以上结果。

考核知识点 8 编写一个 Java 程序，实现一个简单的计算器。该计算器应该能够执行基本的加法、减法、乘法和除法运算。

考核知识点 9 编写一个 Java 程序，针对整数数组 `int[] arr = {1, 5, 2, 4, 3}`，请完成以下操作：

- 1) 找出数组中的最大值和最小值，打印输出。
- 2) 计算数组所有元素的平均值(浮点数)，打印输出。
- 3) 对数组进行排序(升序)，打印排序后的数组。
- 4) 遍历数组，打印出每个元素，格式"元素值为:x"。

考核知识点 10 请叙述大数据、云计算和物联网三者之间的关系。

考核知识点 11 解释矩阵乘法规则，并说明其在数据处理中的应用。

考核知识点 12 关系型数据库与 NoSQL 的区别、优缺点是什么；举例说明 NoSQL 数据库的应用场景，并简要介绍其中一个数据库的特点。

考核知识点 13 请叙述什么是数据挖掘，常用的技术有哪些（不少于 3 种）。

考核知识点 14 请叙述大数据处理需要经过哪些步骤，各步骤主要有什么作用。

考核知识点 15 请叙述常见的大数据框架有哪些，列举其优缺点。