

# 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

本专业大类坚持立德树人根本任务，适应区域经济社会发展需求，立足甘肃、面向西北，培养德、智、体、美、劳全面发展的高素质应用型人才，具有良好政治素养和职业道德，知识结构合理，富有创新精神和实践能力，具备数据预处理、数据存储、数据分析等方面的知识。能够适应未来职业和社会发展，能够数据采集、数据预处理、数据存储、数据分析、数据可视化以及大数据工程运维等工作的专门人才。

### 二、命题的指导思想和原则

（一）命题内容紧密围绕 Python 程序设计、Linux 系统基础、大数据导论等基础课程考核，重点考核学生的理解、掌握和综合运用的能力。

（二）命题以科学性、明确性、全面性、整体性原则为指导，注重考查学生的基础知识、数据可视化、大数据分析及应用能力。问答题目数量共计 15 个，每问难度一致，问答题目表达简练、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（50 分）四部分内容。

### 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电子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导论、程序设计与分析的专业综合能力。

3. 考生需按照公布的测试考核知识点内容及数量进行答辩准备。

4. 考生在测试当日进行现场抽签抽取答辩题号，并按照所抽取的题号进行现场作答。

5. 考生的应试过程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：**在 Linux 中编写 Shell 脚本程序“compare.sh”，当用户从终端输入的数字和预设数字（随机生成一个小于 100 的数字）一样时，直接退出，否则让用户一直输入，并提示用户的数字比预设的数字大或者小。

**考核知识点 2：**使用 Shell 命令完成以下目录和文件的相关操作。

1) 运用 Shell 命令创建一个名为 ‘team’ 的用户组。

2) 新创建一个名为 ‘member’ 的用户，并且把该用户添加到 ‘team’ 用户组中。

3) 在 ‘member’ 用户的主目录下创建一个名为 ‘project’ 的目录，赋予其读写执行权限，然后进入此目录。

4) 在 ‘project’ 目录里创建 ‘doc1.txt’ 和 ‘doc2.txt’ 两个文件。给 ‘doc1.txt’ 赋予读写执行权限，给 ‘doc2.txt’ 赋予只读权限。确保这两个文件的所有者是 ‘member’，所属用户组为 ‘team’。

5) 在 project 目录下创建一个名为 ‘sub\_project’ 的子目录，赋予其读写执行权限后进入该子目录。保证 ‘sub\_project’ 目录的所有者是 ‘member’，所属用户组为 ‘team’。

**考核知识点 3:** 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 4:** 请编写一个 Python 函数，实现从终端输入 2 个正整数，求这两个数的最大公因数，并给出返回值。

**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，格式为"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：**请简述人工智能的定义，并简述机器学习、深度学习，人工智能的区别。

**考核知识点 12：**请叙述常见的大数据框架有哪些，列举其优缺点。

**考核知识点 13：**当前主要的深度学习框架有哪些？分别简述应用领域及其优点？

**考核知识点 14：**什么是 Hadoop？它在大数据处理中的作用是什么？

**考核知识点 15：**简述 Python 中 Pandas 库的主要功能及其在数据分析中的应用。