

数据科学与大数据技术专业中本贯通《技能水平测试》 考试大纲

本考试大纲仅适用于上海市第二轻工业学校 2025 年数据科学与大数据技术专业中本贯通转段考试的学生。专业技能水平考试是中本贯通学生完成中职阶段学习后，进入本科阶段学习前必须参加的转段考试中的一部分。主要目的是考核学生的工程基础理论知识、职业素养、技能操作水平以及解决实际工程问题的知识与技能的综合运用能力。

一、考试科目及分值

- 1.考试科目：理论基础（50%）、实践操作（50%）。
- 2.分值设置：满分 300 分，其中理论基础 150 分、实践操作 150 分。

二、考试形式

1.理论基础

《计算机与数据科学导论》、《Python 基础》、《数据库入门》采用闭卷笔试方式，考试时间为 90 分钟。

题型包括：选择题、判断题、简答题等。

2.实践操作

《Linux 操作系统》、《Python 基础》、《大数据平台安装与部署》、《数据库入门》采用现场机考方式，考试时间为 90 分钟。

题型包括：代码填空题、案例分析题等。

三、考试内容及要求

（一）计算机与数据科学导论（理论基础：30 分）

1.考试目标

考查学生对计算机与数据科学领域基本概念、原理和方法的理解与掌握程度。评估学生运用相关知识分析和解决实际问题的能力。

2.考试内容

（1）数据科学基础，包括数据科学所需知识体系。

（2）数据获取与预处理，包括如何从不同来源获取数据，并进行必要的清洗和预处理以备分析。

(3) 数据探索与可视化，包括通过图表和统计摘要探索数据的重要性，以及如何利用可视化工具传达信息。

(4) 统计学与模型，如统计学方法的分类、自变量与因变量、数据抽样等。

(5) 机器学习与算法，如机器学习中的数据划分、机器学习算法的类型等。

3.考试要求

准确理解并记忆计算机与数据科学的基本术语和概念。能够运用所学原理和方法分析计算机系统运行过程中的问题。掌握基本数据结构的操作和算法设计思路，能够编写简单的程序代码解决实际问题。

(二) Python 基础（理论基础：80 分；实践操作：50 分）

1.理论基础考试目标

检验学生对 Python 语言语法、特性和常用库的掌握情况。评估学生运用 Python 解决数学问题和数据处理任务的能力。

2.理论基础考试内容

(1) Python 语言基础，包括变量、数据类型、运算符、表达式、控制结构（顺序、选择、循环）。

(2) 函数定义与使用，包括函数参数传递、返回值、递归函数等。

(3) 常用数据结构，如字符串、列表的操作和应用。

(4) Python 在数学计算和数据处理方面的应用，如使用 numpy、pandas 等库进行数据操作和分析。

(5) Python 在可视化方面的应用，如使用 Matplotlib 等库进行数据图表绘制和可视化。

3.理论基础考试要求

熟练掌握 Python 语言的基本语法结构，能够编写正确的 Python 程序代码。理解面向对象编程的思想，能够运用类和对象解决实际问题。能够运用 Python 相关库进行数据处理和数学计算，解决简单的实际问题。

4.实践操作考试目标

考查学生实际运用 Python 语言进行程序开发的能力。评估学生在规定时间内完成特定编程任务的能力和代码质量。

5.实践操作考试内容

(1) 根据给定的问题描述，使用 Python 语言编写程序实现特定功能，如数据处理、算法实现、图形绘制等。

(2) 对现有 Python 程序进行调试和优化，查找并修复程序中的错误，提高程序的运行效率。

6.实践操作考试要求

在规定时间内完成编程任务，程序能够正确运行并输出预期结果。代码结构清晰、逻辑合理、注释规范，遵循良好的编程习惯。

(三) Linux 操作系统（实践操作：20 分）

1.考试目标

考查学生对 Linux 操作系统的基本操作和管理能力。评估学生在 Linux 环境下进行系统配置、文件管理和服务管理的实践技能。

2.考试内容

(1) Linux 系统的安装与基本配置，包括分区规划、用户创建、权限设置等。

(2) 文件和目录操作，如文件创建、复制、移动、删除、权限修改，目录创建、切换等。

(3) 文本编辑工具的使用，如 vi/vim 编辑器的基本操作。

(4) 进程管理，包括进程查看、启动、停止、后台运行等操作。

(5) 软件包管理，如使用 yum 等工具进行软件安装、更新和卸载。

(6) 网络配置与服务管理，如 IP 地址配置、网络服务（如 SSH、HTTP 等）的启动和停止。

3.考试要求

熟练掌握 Linux 系统的基本命令，能够准确、高效地完成文件和目录操作。能够根据需求进行系统配置，如用户管理、网络配置等。掌握常用服务的管理和配置方法，确保服务正常运行。

(四) 大数据平台安装与部署（实践操作：50 分）

1.考试目标

考查学生对大数据平台（如 Hadoop、Spark 等）的安装、配置和部署能力。评估学生在实际环境中搭建大数据处理环境的实践技能。

2.考试内容

(1) 大数据平台的架构和组件，如 Hadoop 的 HDFS、MapReduce，Spark 的核心组件等。

(2) 大数据平台的安装过程，包括系统依赖安装、软件包解压、配置文件修改等。

(3) 集群配置与管理，如节点配置、参数调整、服务启动和停止。

(4) 分布式文件系统的使用，如在 HDFS 上进行文件上传、下载、目录操作等。

(5) 简单任务的提交和执行，如使用 MapReduce 提交计算任务。

3.考试要求

能够独立完成大数据平台的安装和基本配置，确保平台正常运行。熟练掌握分布式文件系统的基本操作，能够进行数据存储和管理。能够提交和监控简单的大数据计算任务，验证平台的功能。

(五) 数据库入门（理论基础：40 分；实践操作：30 分）

1.考试目标

考查学生对 MySQL 数据库管理系统的基本操作能力。评估学生在数据库设计、创建、数据操作和查询优化方面的实践技能。

2.考试内容

(1) 数据库系统概述，包括数据库的基本概念、数据模型（如关系模型）、数据库管理系统的功能。

(2) 数据库设计，如根据需求进行实体-关系模型设计，转换为数据库表结构。

(3) SQL 语言基础，包括数据定义语言（DDL）创建数据库、表、视图等，数据操纵语言（DML）进行数据插入、更新、删除和查询操作。

(4) 数据库索引的创建和使用，理解索引对查询性能的影响。

(5) 数据库用户管理和权限设置。

3.考试要求

能够根据需求设计合理的数据库结构，并使用 SQL 语句创建数据库和表。熟练掌握 SQL 语言的数据操作和查询语句，能够准确地进行数据插入、更新、

删除和查询。能够创建合适的索引提高查询性能，进行简单的用户管理和权限设置。

四、参考教材

1. 《数据科学导论(基于 Python 语言)》，[朝乐门]，[人民邮电出版社]，[2021 年 1 月第一版]，[9787115548207]。
2. 《Python 程序编写入门》，[苏东伟]，[高等教育出版社]，[2023 年 7 月第一版]，[978-7-04-058756-2]。
3. 《Linux 网络操作系统项目教程（RHEL 8/CentOS 8）（第 4 版）》，[杨云]，[人民邮电出版社]，[2022 年 1 月第四版]，[978-7-115-56796-3]。
4. 《大数据平台运维（初级）》，[新华三技术有限公司]，[电子工业出版社]，[2020 年 10 月第一版]，[9787121397455]。
5. 《MySQL 数据库基础实例教程》，[周德伟]，[高等教育出版社]，[2023 年 12 月]，[9787040605174]。