

湖南生物机电职业技术学院  
学生专业技能考核标准

**大数据技术专业**

湖南生物机电职业技术学院  
二〇二二年九月

# 目 录

|                   |   |
|-------------------|---|
| 一、专业名称及适用对象 ..... | 1 |
| 二、考核内容 .....      | 1 |
| (一) 专业基本技能 .....  | 1 |
| (二) 岗位核心技能 .....  | 3 |
| 三、评价标准 .....      | 5 |
| 四、抽考方式 .....      | 8 |

# 湖南生物机电职业技术学院大数据技术专业

## 学生专业技能抽查标准

### 一、专业名称及适用对象

#### 1. 专业名称

大数据技术专业（专业代码 610215）。

#### 2. 适用对象

适应大数据技术专业，高等职业院校三年一期全日制在籍学生。

### 二、考核内容

#### （一）专业基本技能

##### 模块一：程序设计

本模块主要测试学生运用 Java 语言进行程序设计的能力。要求学生熟练掌握 Java 语言的基本语法及简单算法的应用，能使用 Java 语言解决常用的程序逻辑问题；能对常用数据结构进行存储和操作。

#### 1. 开发工具使用能力

基本要求：

- （1）能使用常用的开发环境，进行编写、编译、运行 Java 程序；
- （2）能正确调试 Java 程序；
- （3）具有良好的编程习惯。

#### 2. 结构化程序实现能力

基本要求：

- （1）能使用数据类型、变量、常量、运算符、表达式编程；
- （2）能使用数组编程；
- （3）能进行顺序、选择、循环三种控制结构的程序设计；
- （4）能使用函数实现程序功能；
- （5）具有较强的逻辑思维能力。

#### 3. 运用算法设计程序能力

基本要求：

- (1) 能使用查找、排序等常用算法设计程序；
- (2) 具有发散思维和创新意识。
- (3) 具有自主学习和团队协作能力。

## 模块二：数据库设计

本模块以企业事业单位信息管理系统项目开发为背景，完成应用信息系统中数据库开发环境的配置与使用、数据库及数据表的设计、创建与管理、数据表的约束与关系、数据库访问和数据库备份管理等工作内容。基本涵盖了程序员、数据库管理员岗位从事应用系统数据库的设计与开发工作所需的基本技能。

### 1. 数据库管理系统的配置与使用

基本要求：

(1) 能使用 MySQL 数据库管理系统，在 Windows 操作系统上进行相关参数的配置，完成数据库管理系统的安装；

- (2) 能使用 MySQL 数据库管理工具企业管理器；
- (3) 能完成数据库服务器的启动与停止。

### 2. 数据库及数据表的创建与管理

基本要求：

- (1) 能使用企业管理器或 SQL 语句实现数据库的创建和删除；
- (2) 能使用企业管理器或 SQL 语句实现数据表的创建和删除；

### 3. 数据表的操作

基本要求：

- (1) 能使用企业管理器或 SQL 语句实现数据表记录的添加；
- (2) 能使用企业管理器或 SQL 语句实现数据表记录的删除；
- (3) 能使用企业管理器或 SQL 语句实现数据表记录的修改；
- (4) 能使用企业管理器或 SQL 语句实现数据表的基本查询；

### 4. 创建数据表的约束和关系

基本要求：

- (1) 能实现数据表的主、外键约束的添加、修改或删除操作；
- (2) 能实现唯一约束、检查约束、默认约束的添加、修改或删除；

(3) 能根据业务系统需求对数据表的约束及关系进行规范命名, 并具有良好的可读性。

## 5. 数据库的备份与恢复

基本要求:

- (1) 能使用数据库管理工具完成数据库的备份;
- (2) 能使用数据库管理工具完成数据库的恢复。

## (二) 岗位核心技能

### 模块一: 大数据 Hadoop 综合实践

本模块主要测试学生运用 Hadoop 离线处理数据的能力, Hadoop 离线处理目前流行采用 Hive 数据分析技术。要求学生熟练掌握 Hadoop 服务启动, Hive 服务启动, 数据库及数据表的创建, 数据加载及插入, 数据分析与统计, 能使用 Hive 技术解决常用的现实中大数据离线分析和统计问题; 能对分析结果进行大数据存储。

#### 1. 工具使用能力

基本要求:

- (1) 能使用远程登录工具, 远程登录 Hadoop 系统;
- (2) 能使用工具进行 Hive 的本地或者远程操作

#### 2. 大数据离线处理能力

基本要求:

- (1) 掌握 Hive 数据仓库的创建
- (2) 掌握 Hive 数据表的创建
- (3) 掌握 Hive 数据表的加载
- (4) 掌握 Hive 数据表的查询
- (5) 掌握 Hive 数据表的分析统计。
- (6) 掌握 Hive 数据表的导出

#### 3. 运用不同逻辑处理大数据的能力

基本要求:

- (1) 掌握 Hadoop 离线处理的使用场合

(2) 掌握 Hadoop 离线处理 Hive 技术的操作流程

(3) 具有自主学习和团队协作能力。

## 模块二：大数据技术应用综合实践

本模块要求学生能够安装、配置开发环境；能够运用一种基于 Web 的数据采集；能够实现数据预处理；能够运用开发工具对数据进行分析；实现数据可视化。

### 1. 安装与配置开发环境

基本要求：

掌握运行环境的安装与配置。

### 2. 数据采集

熟练掌握通过 Python 完成网页数据采集程序，基本要求：

(1) 掌握新建、保存、运行项目和文件正确。

(2) 掌握认识 HTML 脚本语言，网页抓包分析正确，运行结果进行了截图。

(3) 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。

(4) 掌握将数据写入 csv 文件。

### 3. 数据处理和分析

基本要求：

(1) 掌握复制、打开、保存、运行项目和文件正确。

(2) 掌握 readfile(filename)、printInfo(filename)、insertDate(filename)、aqi(filename)等函数使用。

(3) 掌握模块导入、函数调用正确。

### 4. 数据可视化

掌握在网页中以图表形式对分析结果进行可视化呈现，基本要求：

(1) 掌握新建、保存、运行项目和文件；

(2) 掌握导入 pandas、matplotlib 模块；

(3) 掌握中文显示；

(4) 掌握读取 csv 文件；

(5) 掌握设置画布大小、设置标题、按 AQI 列升序排序、绘制直方图、X

轴标签和 Y 轴标签、保存和显示图片。

### 三、评价标准

1. 评价方式：本专业技能考核采取过程考核与结果考核相结合，技能考核与职业素养考核相结合。根据考生操作的规范性、熟练程度和用时量等因素评价过程成绩；根据设计作品、运行测试结果和提交文档质量等因素评价结果成绩。

2. 分值分配：本专业技能考核满分为 100 分，其中专业技能占 90%，职业素养占 10%。

3. 技能评价要点：根据模块中考核项目的不同，重点考核学生对该项目所必须掌握的技能和要求。虽然不同考试题目的技能侧重点有所不同，但完成任务的工作量和难易程度基本相同。各模块的评价要点和评分细则如表 1 所示，职业素养的评价要点和评分细则如表 2 所示。

表 1 大数据技术专业技能考核评价要点

| 序号 | 类型     | 模块    | 项目             | 评价要点  | 评分细则   |   |   |
|----|--------|-------|----------------|---|--|---|---|
| 1  | 专业基本技能 | 程序设计  | 开发平台的配置与使用     | <ol style="list-style-type: none"> <li>1. 正确配置软件开发环境；</li> <li>2. 开发环境配置过程符合职业规范。</li> <li>3. 具有良好的编程习惯。</li> </ol>   | <ol style="list-style-type: none"> <li>1. 未按要求提交正确格式的源文件，记 0 分。</li> </ol>   |   |   |
|    |        |       | 程序的编写与实现       | <ol style="list-style-type: none"> <li>1. 正确定义变量、常量，名称符合命名规范；</li> <li>2. 正确使用运算符、表达式进行编程；</li> <li>3. 正确使用顺序、分支、循环三种控制结构，实现项目的业务逻辑单元；</li> <li>4. 正确使用数组等基本数据结构进行编程；</li> <li>5. 程序书写结构良好，注释清晰，可维护性好；</li> <li>6. 程序设计合理、语法正确、功能正确完备，并生成可执行文件；</li> <li>7. 开发过程遵循软件开发的规范。</li> <li>8. 具有较强的逻辑思维能力、调试能力和解决问题的能力。</li> </ol> | <ol style="list-style-type: none"> <li>1. 按照任务要求实现相应功能，否则记 0 分。</li> </ol>   |   |   |
|    |        |       | 运用算法设计程序的编写与实现 | <ol style="list-style-type: none"> <li>1. 正确使用冒泡排序算法以及二分法算法设计程序；</li> <li>2. 正确选择合理的算法提升、优化程序效率，具有程序设计发散思维和创新意识；</li> <li>3. 具备扎实的编程语言应用能力、灵活的算法设计能力。</li> </ol>  | <ol style="list-style-type: none"> <li>1. 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。</li> <li>2. 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、使用空行不规范扣 1 分，扣完为止。</li> <li>3. 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。</li> </ol> |   |   |
|    |        | 数据库设计 | 数据库管理系统的配置与使用  | <ol style="list-style-type: none"> <li>1. 正确根据项目的需求选取数据库系统产品；</li> <li>2. 正确使用数据库管理工具，实现系统用户的管理与配置；</li> <li>3. 正确使用数据库管理工具，实现数据库服务器的启动与停止。</li> </ol>  | <ol style="list-style-type: none"> <li>1. 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。</li> </ol>  |   |   |
|    |        |       | 数据库及数据表的创建与管理  | <ol style="list-style-type: none"> <li>1. 正确使用 SQL 语句完成数据库的创建和管理。</li> <li>2. 正确使用 SQL 语句完成数据表的创建和管理。</li> </ol>  | <ol style="list-style-type: none"> <li>1. 没有成功创建数据库，扣 5-8 分。</li> <li>2. 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。</li> </ol>   |   |   |
|    |        |       | 创建数据表的约束和关系    | <ol style="list-style-type: none"> <li>1. 正确使用 SQL 语句实现数据表的主、外键约束的添加、修改或删除操作；</li> <li>2. 正确使用 SQL 语句实现唯一约束、检查约束、默认约束的添加、修改或删除。</li> </ol>  | <ol style="list-style-type: none"> <li>1. 约束创建不成功每一项扣 3-5 分</li> <li>2. 关系创建不符合要求每一项扣 5 分，扣完为止。</li> </ol>   |   |   |
|    |        |       | 数据表的访问         | <ol style="list-style-type: none"> <li>1. 正确使用 SQL 语句执行数据表记录的增、删、改操作；</li> <li>2. 正确使用 SQL 语句实现数据表的基本查询。</li> </ol>   | <ol style="list-style-type: none"> <li>1. 没有正确写出 SQL 语句每一项扣 4-5 分，扣完为止。</li> </ol>   |   |   |
|    |        |       | 数据库的分离与附加      | <ol style="list-style-type: none"> <li>1. 正确使用数据库管理工具实现数据库的分离；</li> <li>2. 正确使用数据库管理工具实现数据库的附加。</li> </ol>  | <ol style="list-style-type: none"> <li>1. 数据库名命名不规范扣 2 分。数据表命名不规范每张表扣 1 分，扣完为止。</li> <li>2. 字段名命名不规范每项扣 0.5 分，扣完为止。</li> </ol>   |   |   |
|    |        | 2     | 岗位核            | 大数据   | Hive 服务的启动   | <ol style="list-style-type: none"> <li>1. 启动 Hadoop 大数据平台</li> <li>2. 启动 Hive 服务</li> </ol> | <ol style="list-style-type: none"> <li>1. 未启动集群，扣 5 分。</li> <li>2. 未启动 Hive 服务，扣 5 分</li> </ol> |
|    |        |       |                | Had   | Hive 数据库   | <ol style="list-style-type: none"> <li>1. 实现数据仓库创建</li> </ol>                               | <ol style="list-style-type: none"> <li>1. 未创建数据仓库，扣 10 分。</li> </ol>                            |



|             |                 |           |   |   |
|-------------|-----------------|-----------|---|---|
| 心<br>技<br>能 | oop<br>综合<br>应用 | 及数据表创建    | 2. 实现数据表创建  | 2. 未创建数据表, 扣 10 分。  |
|             |                 | 数据表数据加载   | 1. 将各种数据加载到数据表中   | 1. 未正确加载数据表, 扣 10 分   |
|             |                 | 简单数据分析与处理 | 根据数据指标, 进行简单数据分析和处理   | 1. 未能根据数据指标, 进行简单数据分析和处理, 扣 20 分  |
|             |                 | 复杂数据分析处理  | 1. 根据数据指标, 进行复杂的数据分析和处理   | 1. 未能根据数据指标, 进行复杂数据分析和处理, 扣 30 分  |
|             |                 | 数据结果导出    | 根据用户需求, 能够将结果进行数据导出和存储  | 未能将结果进行数据导出和存储, 扣 10 分  |
|             | 大数据技术<br>应用综合实  | 数据采集      | 1. 正确新建、保存、运行项目和文件正确。<br>2. 正确认识 HTML 脚本语言, 网页抓包分析正确, 运行结果进行了截图。<br>3. 正确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。<br>4. 正确将数据写入 csv 文件。<br>5. 开发环境配置过程中, 具备规范化操作与生产意识。                      | 1. 新建、保存、运行项目和文件正确, 扣 5 分。<br>2. 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 2 分。<br>3. 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。<br>4. 将数据写入 csv 文件。代码不正确 1 处扣 1 分。   |
|             |                 | 数据结果导出    | 1. 根据用户需求, 能够将结果进行数据导出和存储   | 1. 未正确使用 HTML 标签与 DIV+CSS 样式完成静态页面的设计与布局, 每一项扣 2 分;<br>2. 未正确使用文本元素、多媒体元素、链接制作网页, 每一项扣 2 分;<br>3. 未使用表格、框架、表单布置网页, 每一项扣 2 分;<br>4. 未在页面设计与实现的过程中, 遵守 Web 前端设计的技术标准, 页面美观大方, 用户体验良好, 每一项扣 2 分。   |
|             |                 | 数据处理和分析   | 1. 正确复制、打开、保存、运行项目和文件正确。<br>2. readFile(filename)函数定义正确。<br>3. printInfo(filename)函数定义正确。<br>4. insertDate(filename)函数定义正确。<br>5. maxCount(filename)函数定义正确。<br>6. 模块导入、函数调用正确。                | 1. 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。<br>2. readFile(filename)函数定义正确。代码不正确 1 处扣 1 分。<br>3. printInfo(filename)函数定义正确。代码不正确 1 处扣 1 分。<br>4. insertDate(filename)函数定义正确。代码不正确 1 处扣 1 分。<br>5. dropData(filename)函数定义正确。代码不正确 1 处扣 1 分。<br>6. 模块导入、函数调用正确。代码不正确 1 处扣 1 分。 |
|             |                 | 数据可视化     | 1. 正确新建、保存、运行项目和文件正确。<br>2. 导入 pandas、matplotlib 模块正确<br>3. 中文显示正确。<br>4. 读取 csv 文件正确。<br>5. 设置画布大小合适。<br>6. 设置标题正确。<br>7. 按人次列降序排序正确。<br>8. 绘制折线图正确。<br>9. X 轴标签和 Y 轴标签正确。<br>10. 保存和显示图片正确。 | 1. 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 2 分。<br>3. 导入 pyecharts 模块及函数正确, 不正确扣 2 分。<br>3. 导入 pandas 模块正确, 不正确扣 1 分。<br>4. 读取 csv 文件中的监测点、AQI 数据正确, 不正确扣 2 分。<br>5. Geo 函数调用正确, 不正确扣 1 分。  |

表 2 各模块职业素养评价

| 评价项目   | 评价要点   | 评分细则                           |
|--------|--|--------------------------------|
| 工作前准备  | 1. 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。                                    | 1. 不进行检查操作扣 5 分。               |
| 专业素养   | 1. 代码符合软件开发规范，命名规范，能做到见名知意；<br>2. 缩进统一，方便阅读；<br>3. 注释规范正确。                 | 1. 至少要有 5 行注释，少 1 行或错 1 行扣 1 分 |
| 职业行为规范 | 1. 着装干净、整洁。<br>2. 举止文明，遵守考场纪律，按顺序进出考场。<br>3. 考试过程中及结束后，考试桌面及地面要符合 6S 基本要求。 | 1. 不符合 6S 基本要求的扣 3-5 分。        |

#### 四、抽考方式

技能考核为现场操作考核，成绩评定采用过程考核与结果考核相结合。具体方式如下：

1. 学校参考模块选取：采用“2+2”的模块选考方式，专业基本技能模块分为程序设计模块和数据库设计模块，岗位核心技能模块分为大数据 Hadoop 综合模块和大数据技术应用综合实践模块。

2. 学生参考模块确定：参考学生按规定比例随机抽取考试模块，其中，50%考生参考专业基本技能模块，50%考生参考岗位核心技能模块。各模块考生人数按四舍五入计算，剩余的尾数考生随机在两类模块中抽取应试模块。

3. 试题抽取方式：学生在两类四个模块题库中随机抽取 1 道试题考核。

表 6 考核模式

| 类型     | 模块              | 考核项目              | 考核要求         | 参考学生 |
|--------|-----------------|-------------------|--------------|------|
| 专业基本技能 | 程序设计模块          | 开发平台的配置与使用        | 三个项目必选<br>必考 | 25%  |
|        |                 | 项目的设计与建模          |              |      |
|        |                 | 程序的编写与实现          |              |      |
|        | 数据库设计模块         | 数据库管理系统的配置与使用     | 五个项目必选<br>必考 | 25%  |
|        |                 | 数据库及数据表的创建与管理     |              |      |
|        |                 | 创建数据表的约束和关系       |              |      |
| 数据表的访问 |                 |                   |              |      |
| 岗位核心技能 | 大数据 Hadoop 综合模块 | Hive 数据仓库创建和创建数据表 | 五个项目必选<br>必考 | 25%  |
|        |                 | 数据表的加载            |              |      |
|        |                 | 数据表的简单分析与处理       |              |      |

|  |               |              |              |     |
|--|---------------|--------------|--------------|-----|
|  |               | 数据表的复杂分析与处理  |              |     |
|  |               | 分析统计结果的导出与处理 |              |     |
|  | 大数据技术应用综合实践模块 | 数据采集         | 三个项目必选<br>必考 | 25% |
|  |               | 数据处理和分析      |              |     |
|  |               | 数据可视化        |              |     |