

# 目 录

<b>一、专业名称及适用对象 .....</b>	<b>1</b>
1. 专业名称.....	1
2. 适应对象.....	1
<b>二、考核目标 .....</b>	<b>1</b>
<b>三、考核内容 .....</b>	<b>1</b>
模块一 Python 程序设计 .....	1
1.技能要求.....	2
2.素养要求.....	2
3.完成时间: .....	2
模块二 Hadoop 平台部署及应用 .....	2
1.技能要求.....	2
2.素养要求.....	3
3.完成时间: .....	3
模块三 数据采集.....	3
1.技能要求.....	3
2.素养要求.....	3
3.完成时间: .....	3
模块四 数据可视化与分析.....	4
1.技能要求.....	4
2.素养要求.....	4
3.完成时间: .....	4
<b>四、评价标准 .....</b>	<b>4</b>
1.评分方式.....	4
2.分值分配.....	4
3.技能评价要点.....	5
<b>五、抽考方式 .....</b>	<b>6</b>
<b>六、附录 .....</b>	<b>6</b>
(一) 相关法律法规(摘录) .....	6
(二) 相关规范与标准(摘录) .....	6

# 湖南商务职业技术学院学生专业技能考核标准

## 一、专业名称及适用对象

### 1. 专业名称

大数据技术与应用（专业代码：610215）。

### 2. 适应对象

高职大数据技术与应用专业全日制在籍大三学生。

## 二、考核目标

依据本专业人才培养方案，本专业技能考核通过设置Python程序设计、数据采集、数据可视化与分析、Hadoop平台部署及应用等4个技能模块，测试学生掌握计算机软硬件基本理论、大数据平台搭建知识、大数据平台运维及测试知识、数据仓库基础知识、数据处理基础知识情况；测试学生具备大数据平台搭建与运维、数据采集、数据加工、数据分析、数据可视化等专业实践能力情况；测试学生具备认知能力、合作能力、创新能力、职业能力等支撑终身发展、适应时代要求的关键能力情况，测试学生具有从事大数据相关岗位的职业素养情况。促进专业不断完善教学基本条件，深化教育教学改革，强化实践教学环节，增强学生创新创业能力，促进学生个性化发展，提升专业建设水平，提升课程教学的有效性，培养适应信息时代发展需要的，符合大数据技术与应用专业设计、开发、操作、维护、营销、服务需要的复合型、创新型和高素质技术技能人才。

在技能抽查考核过程中涉及到的课程有《Python程序设计》、《Hadoop实用技术》、《网络爬虫技术》、《Spark大数据技术与应用》、《大数据ETL技术》、《大数据系统维护与管理》、《数据分析挖掘工具》、《数据可视化技术》等。

## 三、考核内容

根据大数据技术与应用专业培养目标，针对其就面向的就业岗位和能力要求，由浅入深，设计出4个模块进行考核，分别是python程序设计、Hadoop平台部署及应用、数据采集、数据可视化与分析。

表1 考核模块与典型工作任务表

序号	模块名称	典型工作任务
1	Python程序设计	大数据平台开发
2	Hadoop平台部署及应用	大数据运维、大数据架构设计
3	数据采集	大数据的采集
4	数据可视化与分析	数据可视化展示、数据分析

### 模块一 Python程序设计

“Python程序设计”是大数据从业人员必须具备的基本技能，主要用于大数据平台的各

项开发。本模块的主要考核内容为：Python基础入门、函数、面向对象编程、实用模块等。

## 1.技能要求

### (1) Python程序设计基本要求

- ① 掌握 Python 中的常用操作符；
- ② 掌握 Python 的固定语法；
- ③ 掌握列表、元组、字典和集合的概念和基本操作；
- ④ 掌握循环语句、终止语句、通过语句；
- ⑤ 熟练掌握组合语句的编写和应用；
- ⑥ 掌握多变量迭代和列表解析的用法。

### (2) Python函数基本要求

- ① 熟练掌握函数的调用和实现；
- ② 熟练掌握函数嵌套和自定义函数；
- ③ 熟练掌握函数对数据进行处理。

### (3) Python面向对象编程基本要求

- ① 掌握面向对象编程；
- ② 熟练掌握类和对象的概念、迭代器的运用；
- ③ 熟练掌握继承的概念和运用。

### (4) 文件读写基本要求

- ① 熟练掌握文件的读取与写入；
- ② 熟练掌握使用 OS 模块进行工作路径、文件夹等管理的操作。

## 2.素养要求

- ① 具有较强的分析问题的能力、发散思维和创新意识；
- ② 具有良好的编程习惯、较强的逻辑思维能力及综合运用知识的能力；
- ③ 能够完成真实业务逻辑向代码转换；
- ④ 能够独立解决技术问题；
- ⑤ 具备程序员的严谨认真、规范的工作态度和正确的价值观。

## 3.完成时间：

2小时。

## 模块二 Hadoop平台部署及应用

本模块以Hadoop实用技术为主，主要培养学生的大数据运维能力和大数据架构设计能力，主要考核内容包含搭建Hadoop完全分布式集群，HDFS的原理和基础操作，Hive基本应用等。

## 1.技能要求

### (1) Hadoop环境搭建基本要求

- ① 能够配置和管理集群环境，完成相关设备的维护；
- ② 能够完成操作系统的安装和配置；
- ③ 能够完成大数据平台的日常维护；
- ④ 能够实现配置 Hadoop 生态组件的配置。

## (2) 分布式文件系统的使用基本要求

- ① 能根据业务需求，正确设计 HDFS 文件系统；
- ② 能使用 HDFS 文件系统完成项目任务。

## (3) 分布式数据库基本要求

- ① 能正确的根据需求完成常见分布式数据库 Hive 的使用；
- ② 能根据需求完成数据添加、查询操作；
- ③ 具备编写 HQL 语句的能力，能完成简单的数据统计和分析；
- ④ 能够向数据库中提供高质量的数据，完成数据预处理工作。

## 2.素养要求

- ① 能够独立完成环境搭建；
- ② 能够独立解决技术问题；
- ③ 能够规范地操作命令；
- ④ 熟悉相关规范和标准，能识读相关软件技术文档。

## 3.完成时间：

2小时

## 模块三 数据采集

本模块主要培养学生的数据采集能力，考核项目为网络爬虫技术，主要考核内容为：网络爬虫基本原理、网页爬虫的基本方法、网页的分析方法以及数据的储存方法等。

## 1.技能要求

### (1) 网络爬虫技术概述基本要求

- ① 掌握 Python 爬虫的环境配置方法；
- ② 熟悉 HTTP 请求方式与过程；
- ③ 熟悉常见 HTTP 状态码；
- ④ 熟悉 Cookie。

### (2) 网页爬虫基本要求

- ① 掌握使用 urllib3 库、Requests 库实现 HTTP 请求的发送；
- ② 掌握使用 Chrome 开发者工具、正则表达式、Xpath 和 BeautifulSoup 等工具进行解析网页；
- ③ 掌握使用 JSON 库进行数据存储。

## 2.素养要求

- ① 能够完成真实业务逻辑向代码转换；
- ② 能够独立解决技术问题；
- ③ 自学能力强，能够快速准确地查找参考资料；
- ④ 能够按照规范编写技术文档；
- ⑤ 在进行网络爬虫的过程中能注意内容的合法性（不能非法爬取隐私信息等）、规范性。

## 3.完成时间：

2小时

## 模块四 数据可视化与分析

本模块主要培养学生对处理好的数据进行可视化展示和分析的能力，考核内容包括：饼图、柱状图等常用图形可视化，图像参数调整，趋势分析可视化，相关分析可视化，多指标数据可视化，图层叠加，图像组合，图像保存等，然后对已经整理好的数据进行分析 and 给出有价值的意见或建议。

### 1.技能要求

#### (1) 数据可视化技术基本要求

- ① 掌握 matplotlib、pyplot 常用的绘图参数的调节方法；
- ② 掌握图形组合、子图绘制的方法；
- ③ 掌握绘制图形的保存与展示方法；
- ④ 掌握常用饼图、柱状图等图的作用与绘制方法。

#### (2) 分析结果可视化要求

- ① 掌握多指标数据可视化的方法；
- ② 掌握趋势分析可视化的方法；
- ③ 掌握相关分析可视化的方法。

#### (3) 数据可视化分析基本要求

- ① 掌握数理统计分析等分析方法，并能初步进行市场预测；
- ② 掌握数据分析报告撰写方法；
- ③ 掌握趋势分析可视化的方法；
- ④ 能够使用专业知识分析企业数据中展示出来的企业运营问题；
- ⑤ 能够应用现代计算工具和软件分析数据资料。

### 2.素养要求

- ① 具备良好的自学能力，关注互联网新知识、新技术，有一定创新意识；
- ② 理解业务背景，能够完成真实业务逻辑向代码转换；
- ③ 具有一定的视觉审美能力、创新设计理念；
- ④ 具有高度的敬业精神及工作激情，工作态度积极乐观。

### 3.完成时间：

2小时

## 四、评价标准

### 1.评分方式

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85分，职业素养占15分。

### 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。各考核内容的技能评价要点内容如表2所示：

表2 大数据技术专业技能考核评价要点

序号	考核内容	评价要点
1	职业素养	答题文件夹命名规范 各文件的命名规范 代码规范扣 变量名命名规范 考场礼仪规范
2	Python程序设计	Python常用数据类型使用正确，代码正确运行 Python数据结构理解正确，操作正确 Python数据控制流操作正确，选择正确，思路正确 自定义函数正确编写，调用成功 面对对象概念理解正确，类编写正确，实例化成功 文件写入成功，读取成功
3	Hadoop平台部署及应用	Hadoop核心组合理解正确，选择正确 虚拟机安装配置正确 JDK安装正确 Hadoop完全分布式集群环境搭建正确 Hadoop集群的基本信息调用正确 HDFS基本操作编写正确，操作正确 Hadoop相关组件Hive等正确使用 Hadoop组件的应用理解正确
4	数据采集	使用urllib3库、Requests库发送HTTP请求成功，选择的请求方法正确 使用BeautifulSoup分析网页信息正确 使用正则表达式提取网页信息正确 使用XPath提取网页信息正确
5	数据可视化与分析	指定图表类型（饼图、柱状图等）绘制正确 图形组合合理，绘制子图方法正确 趋势分析选择图形合理 相关分析选择图形合理 按照要求进行了基本处理，如添加文字、边框制作等，内容准确、设计美观 背景分析全面、具体、正确 数据分析目标结合公司实际情况，清晰具体有针对性，要有数据表示，有图表说明 数据分析对象应准确、直观、符合逻辑 数据分析选择要有针对性，符合企业实际情况与需要 数据分析方案设计创意好，图文结合、美观、直观，与客户群体、主营业务等相契合

## 五、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. 学生抽取：从高职大数据技术与应用专业全日制在籍大三学生随机抽取20%参加技能抽查考试（最少20人）。学生按应考人数1: 1.1比例抽取。确定抽查专业后，由现场考评组长或考评人员从教务处提供的抽查全名单中随机抽取学生。
2. 模块抽签：测试前，由现场考评组长或考评员根据专业技能抽查考试标准的模块数量确定不同模块考生的数量，学生通过抽签确定参考哪一个模块。
3. 试题抽签：在学生参考模块确定以后，由现场考评组长或考评员从考生对应模块的已封存好的试题中抽取1道试题作为考生的测试题。
4. 工位抽签：参加测试的学生须要在测试前达到对应的考场地点，考评员组织学生随机抽签确定工位号，并登记备案。

## 六、附录

### （一）相关法律法规（摘录）

1.中华人民共和国网络安全法（中华人民共和国主席令 第五十三号）

第二十一条 国家实行网络安全等级保护制度。网络运营者应当按照网络安全等级保护制度的要求，履行下列安全保护义务，保障网络免受干扰、破坏或者未经授权的访问，防止网络数据泄露或者被窃取、篡改：

（一）制定内部安全管理制度和操作规程，确定网络安全负责人，落实网络安全保护责任；

（二）采取防范计算机病毒和网络攻击、网络侵入等危害网络安全行为的技术措施；

（三）采取监测、记录网络运行状态、网络安全事件的技术措施，并按照规定留存相关的网络日志不少于六个月；

（四）采取数据分类、重要数据备份和加密等措施；

（五）法律、行政法规规定的其他义务。

2.国务院关于印发促进大数据发展行动纲要的通知 国发[2015]50号

（三十）提升产业支撑能力。进一步健全创新体系，鼓励相关企业、高校和科研机构开展产学研合作，推进大数据协同融合创新，加快突破大规模数据仓库、非关系型数据库、数据挖掘、数据智能分析、数据可视化等大数据关键共性技术，支持高性能计算机、存储设备、网络设备、智能终端和大型通用数据库软件等产品创新。支持企事业单位开展大数据公共技术服务平台建设。鼓励具有自主知识产权和技术创新能力的大数据企业做强做大。推动各领域大数据创新应用，提升社会治理、公共服务和科学决策水平，培育新的增长点。落实和完善支持大数据产业发展的财税、金融、产业、人才等政策，推动大数据产业加快发展。

3.《计算机软件著作权登记办法》2002年

4.《计算机软件保护条例》2001年

5.《计算机病毒防治管理办法》2000年

6.《计算机信息系统安全专用产品检测和销售许可证管理办法》1997年

7.《中华人民共和国计算机信息系统安全保护条例》1994年

### （二）相关规范与标准（摘录）

本专业标准主要依据的大数据行业相关规范与标准如表3所示。

表3 引用技术标准和规范

序号	中文标准名称	引用标准来源或链接网址
1	大数据标准化 白皮书 (2018)	<a href="http://www.cesi.cn/201803/3709.html">http://www.cesi.cn/201803/3709.html</a>
2	GB/T 35273- 2017 《信息 安全技术 个 人信息安全规 范》	<a href="https://www.tc260.org.cn/front/postDetail.html?id=20180124211617">https://www.tc260.org.cn/front/postDetail.html?id=20180124211617</a>
3	《信息安全技 术 大数据服 务安全能力要 求》	<a href="https://www.tc260.org.cn/front/postDetail.html?id=20161221102931">https://www.tc260.org.cn/front/postDetail.html?id=20161221102931</a>
4	《信息安全技 术 个人信息 安全规范（草 案）》	<a href="https://www.tc260.org.cn/front/postDetail.html?id=20190201173320">https://www.tc260.org.cn/front/postDetail.html?id=20190201173320</a>