



山西职业技术学院

SHANXI POLYTECHNIC COLLEGE

# 大数据技术与应用专业 人才培养方案

(2019 级)

二〇一九年六月



# 目 录

一、专业名称及代码 .....	1
二、招生对象 .....	1
三、修业年限 .....	1
四、职业面向 .....	1
五、培养目标与培养规格 .....	1
六、课程设置 .....	4
七、学时分配 .....	6
八、教学进程总体安排 .....	8
九、毕业标准 .....	10
十、实施保障 .....	10
附件 1 大数据技术与应用专业人才需求调研报告 .....	20
附件 2 大数据技术与应用职业标准 .....	31
附件 3 大数据技术与应用专业核心课程标准 .....	33
《Java 程序设计(一)》课程标准 .....	33
《Java 程序设计(二)》课程标准 .....	38
《MySQL 数据库》课程标准 .....	43
《Python 程序设计》课程标准 .....	47
《Hadoop 系统基础》课程标准 .....	53
《HBase 系统基础》课程标准 .....	59
《Hive 数据仓库系统》课程标准 .....	63
《Spark 大数据技术与应用》课程标准 .....	68



## 一、专业名称及代码

专业名称：大数据技术与应用

专业代码：610215

## 二、招生对象

高中阶段教育毕业生或具有同等学力者。

## 三、修业年限

高等职业学校学历教育修业年限为3年。

## 四、职业面向

所属专业 大类	所属专业类	对应行业	主要职业 类别	主要岗位类别(或技术领域)	职业资格证书或技能等级 证书
电子信息 大类	计算机类	软件和信息技术服务业(65)	软件和信息技术服务人员	1. 大数据应用开发 2. 大数据系统运维	大数据运维工程师 大数据实施工程师

## 五、培养目标与培养规格

### (一) 培养目标

培养思想政治坚定、德技并修、全面发展，适应山西区域经济发展需要，具有良好职业素养，掌握大数据基本理论、方法与技能，面向大数据应用企事业单位，能够从事使用大数据技术，从事大数据系统搭建与运维、大数据获取与存储、大数据处理与分析等工作的高素质技术技能人才。

### (二) 培养规格

#### 1. 素质

- (1) 热爱祖国、遵纪守法，树立科学的世界观、人生观和价值观，具有良好的思想品德、社会公德和坚定的政治素养；
- (2) 具有一定的科学素养和文学、艺术修养；
- (3) 具有积极健康、乐观向上的身心素质；
- (4) 具有爱岗、敬业、奉献、协作等职业素养；
- (5) 具有诚信品格、服务意识、质量意识；
- (6) 具有良好的职业道德与职业操守，具备较强的组织观念和集体意识；
- (7) 具有一定的数理与逻辑思维；
- (8) 具有一定阅读并正确理解分析报告和项目建设方案的能力；

- (9) 具有阅读本专业相关中英文技术文献、分析与处理，进行文档管理的能力；
- (10) 具有一定的工程意识和效益意识；
- (11) 具有创新精神和开拓能力。

## 2.知识

- (1) 掌握必备的体育健身基础知识和相关心理健康知识；
- (2) 掌握必备的计算机应用、英语、数学的基本知识；
- (3) 掌握公共安全、自身安全防范的基本知识；
- (4) 了解相关法律法规及国际通用惯例；
- (5) 掌握计算机软硬件基础知识；
- (6) 掌握程序设计、软件需求分析基础知识；
- (7) 掌握数据库原理、操作系统原理基础知识；
- (8) 具备网站设计、计算机网络基础知识；
- (9) 了解软件部署、软件系统安装运维知识；
- (10) 掌握大数据系统工作原理及其搭建维护知识；
- (11) 了解云计算平台工作原理及基本操作知识；
- (12) 掌握数据获取和预处理知识；
- (13) 掌握 Map Reduce 编程知识；
- (14) 掌握数据仓库知识；
- (15) 掌握数据处理基本知识；
- (16) 掌握数据分析基本知识；
- (17) 掌握数据挖掘基本知识；
- (18) 了解 IT 领域前沿动态，掌握部分相关新知识；
- (19) 具备一定的经营管理知识、经济效益意识及成本核算知识。

## 3.能力

### 职业基本能力：

- (1) 具有运用辩证唯物主义的基本观点及方法认识、分析、解决问题的能力；
- (2) 具有一定的应用文写作、英语听说读写及数学运用能力；
- (3) 具有计算机应用系统的操作和组装、维护与维修能力；
- (3) 具有数据库系统的基本操作能力；
- (4) 具有基本的程序设计能力；

- (5) 具有网站页面设计与制作的能力；
- (6) 具有网站后台程序设计和网络数据库设计能力，网络应用开发技术；
- (7) 具有常用办公软件、工具软件的使用能力，利用 Office 工具进行项目开发文档的整理（Word）、报告的演示（PowerPoint）、表格的绘制与数据的处理（Excel），利用 Visio 绘制流程图；
- (8) 具有阅读并正确理解需求分析报告和项目建设方案的能力；
- (9) 具有阅读本专业相关中英文技术文献、资料的能力；
- (10) 具有熟练查阅各种资料，并加以整理、分析与处理，进行文档管理的能力；
- (11) 具有通过系统帮助、网络搜索、专业书籍等途径获取专业技术帮助的能力；
- (12) 具有具有解决实际问题的综合技术能力。

**专业核心能力：**

- (1) 具有各种大数据系统安装、维护与维修能力，包括 Hadoop、HBase、Hive、Spark 等系统；
- (2) 具有使用各种大数据系统解决实际问题的能力，包括 Hadoop、HBase、Hive、Spark 等系统；
- (3) 具有获取数据和对数据进行预处理的能力；
- (4) 具有大数据应用开发能力，包含批处理、实时流式处理、随机查询、数据挖掘分析等模式；
- (5) 基本的大数据分析和数据挖掘能力。

**其他能力：**

- (1) 方法能力：分析问题与解决问题的能力、应用知识能力、创新能力；
- (2) 工程实践能力：人员管理、时间管理、技术管理、流程管理等能力；
- (3) 组织管理能力。

## 六、课程设置

### (一) 课程结构

公共基础课 (13 门)	专业课 (14 门)	专业拓展课 (5 门)
国防教育与军事训练、入学教育	Java 程序设计(一) ★	计算机组装与维护
思想道德修养与法律基础	HTML 5 网页设计	计算机网络基础
毛泽东思想和中国特色社会主义理论体系概论	Linux 操作系统	大数据导论
大学语文	Java 程序设计(二) ★	Shell 编程
应用数学	MySQL 数据库★	专业英语
基础英语	Python 程序设计★	
体育	Hadoop 系统基础★	
形势与政策	HBase 系统基础★	
心理健康	Hive 数据仓库系统★	
安全教育	Spark 大数据技术与应用★	
计算机应用基础	大数据可视化技术	
大学生职业发展与就业指导	毕业设计 (论文)	
创新创业教育	跟岗实习	
	顶岗实习	

备注：标注“★”的课程为专业核心课程

### (二) 专业核心课程简介

课程名称	Java 程序设计(一)	开设学期	第一学期		
课程代码		参考学时	56	学分	4

本课程主要讲解 Java 变量和数据类型，选择结构，循环结构，流程控制重难点串讲，多重循环及程序调试，循环进阶重难点串讲，一维数据及经典应用等内容。本课程吸收了最新的 Java 技术和语法，融合了编程思想，理论，实践，应用于一体；并且知识点前呼后应，通过课堂教学和实践教学相结合，使学生能够深入理解面向对象概念，清楚的了解 Java 软件开发流程，建立起应用程序的概念，最终能够掌握 Java 软件开发的基本方法，基本技能的能力。

课程名称	Java 程序设计(二)	开设学期	第二学期		
课程代码		参考学时	72	学分	4



本课程分为两部分，Java 面向对象与 JDK 核心 API，分别讲解了面向对象特性(封装，继承，多态等)，语言核心包，异常处理，常用工具包，集合框架等。是学生可以掌握面向对象的基本原则以及在编程实践中的意义，掌握 Java 面向对象编程基本实现原理。熟练掌握 JDK 核心 API 编程技术，理解 API 设计原则，具备熟练的阅读 API 文档的能力，为后续的课程学习打下坚实的语言基础。

课程名称	MySQL 数据库	开设学期	第三学期		
课程代码		参考学时	72	学分	4

MySQL 数据库是当前最流行的关系型数据库之一，本课程以 MySQL 数据库为核心，主要授课内容包括数据库安装，数据库表创建管理，使用 SQL 语句实现数据添加，修改，查询以及事务，视图，索引，备份和恢复等知识，为保证最优学习效果，紧密结合实际使用，利用大量案例说明和实践，让学生都能完成一个典型系统的数据库的设计。

课程名称	Python 程序设计	开设学期	第三学期		
课程代码		参考学时	72	学分	4

本课程主要学习 Python 数据类型，控制结构，正则表达式等基本知识以及大数据处理模块等内容，学习 Python 环境搭建及 Python 数据类型，Python 基础及函数，Python I/O 及文件读写，Python 面向对象编程，Python 库介绍（Numpy，Pandas），Python 数据爬取，Python 数据采集，整理及清洗等方面的应用，使学生通过 Python 编程语言的爬虫抓取工具，实现对网络数据抓取，以及使用 Python 的 Scrapy 框架来进行爬虫的抓取。通过 Python API 接口，能对计算分析结果导出到 HDFS 中，并提供报表，日志分析等结果；通过 Python 实现数据的展现，使学生具备使用 Python 对大数据进行操作及综合分析的能力

课程名称	Hadoop 系统基础	开设学期	第三学期		
课程代码		参考学时	108	学分	6

本课程主要学习 Hadoop 的体系架构，原理及搭建方法，HDFS 文件系统架构，原理及使用方法，MapReduce 计算框架的工作流程，基于 MapReduce 的大数据应用开发，使学生理解 Hadoop 在大数据技术中的地位，理解 Hadoop 的体系结构和工作原理，理解 HDFS 文件系统结构和工作原理，理解 MapReduce 计算框架的工作流程，具有搭建和配置 Hadoop 系统的能力，具有基于 MapReduce 编程的能力及使用 HDFS 文件系统的能力

课程名称	HBase 系统基础	开设学期	第四学期		
课程代码		参考学时	68	学分	4

本课程主要学习 HBase 的体系架构及实现方法，逻辑模型及物理模型，HBase 的搭建方法和使用方法，HBase 表的设计，创建和使用，使学生了解 HBase 在大数据技术中的地位，理解 HBase 和关系型数据库的本质差别，掌握 HBase 的体系结构和工作原理，具备搭建和配置 HBase 系统的能力，具备设计，创建和使用 HBase 表的能力

课程名称	Hive 数据仓库系统	开设学期	第四学期		
课程代码		参考学时	68	学分	4

本课程主要学习数据仓库的基本概念和作用，学习 Hive 的体系架构及搭建方法以及 Hive DDL，DML 和 SQL 操作方法，HiveQL 使用实例，使用 Hive 建立数据仓库的过程等，使学生理解 Hive 在大数据技术中的地位，理解数据仓库与数据库的本质区别，掌握 Hive 的工作原理，使学生具备定义，创建和使用数据仓库的能力，搭建和配置 Hive 系统的能力，存储 Hive 元数据的能力

课程名称	Spark 大数据技术与应用	开设学期	第四学期		
------	----------------	------	------	--	--

课程代码		参考学时	102	学分	6
------	--	------	-----	----	---

本课程主要学习 spark 的基本概念和作用, spark 使用 Scala 语言进行实现, 它是一种面向对象, 函数式编程语言, 能够像操作本地集合对象一样轻松的操作分布式数据集。本课程学习 Scala 基础, Scala 高级使用, Apache Spark 基础及架构, Apache Spark 分布式计算原理, Spark SQL 精华与 Hive 的集成, 基于 Spark GraphX 的图形数据分析等内容。让学生对大数据拥有数据清理, 分析的能力。

## 七、学时分配

表 7-1 教学活动按周分配表

学期	入学教育及军训	课堂教学	集中实训	教学周合计	机动	考试周	学期小计	假期	总计
1	2	14		16	1	1	18	6	24
2		18		18	1	1	20	6	26
3		18		18	1	1	20	6	26
4		17	1	18	1	1	20	6	26
5		0	18	18	1	1	20	6	26
6		0	20	20	0	0	20		20
总计	2	67	39	108	5	5	118	30	148

表 7-2 学期教学任务书

学期	课程代码	课程名称	课程类型	教学周数	建议周学时	学时数 (理论+实践)
第一 学期	2100001	国防教育与军事训练、入学教育	C	2	√	0+48
	1200009	思想道德修养与法律基础	A	14	2	20+7
	1200026	形势与政策	A	√	√	8+0
	1200012	心理健康	A	14	1	14+0
	2100003	安全教育	A	√	√	4+0
	1200030	大学生职业发展与就业指导	A	14	1	16+5
	1110046	大学语文	A	14	2	30+0
	1110044	应用数学	A	14	4	48+8
	1110049	基础英语	A	14	2	28+0
	1400007	体育	B	14	2	4+24
	0911001	计算机应用基础	B	14	4	12+44
	0511001	Java 程序设计（一）	B	14	4	28+28
	0911009	计算机组装与维护	B	14	2	18+10
	0917004	HTML5 网页设计	B	14	2	15+15
	合计学时					<b>245+189=434</b>
第二	1200010	思想道德修养与法律基础	A	18	2	20+7
	1200027	形势与政策	A	√	√	8+0

## 大数据技术与应用专业人才培养方案

学期	2100004	安全教育	A	√	√	4+0
	1110058	大学语文	A	18	2	30+0
	1110045	应用数学	A	18	2	24+4
	1110050	基础英语	A	18	4	56+0
	1400008	体育	B	18	2	4+24
	2100009	创新创业教育	B	18	2	12+20
	0911008	计算机网络基础	B	18	2	18+18
	0917002	Java 程序设计（二）	B	18	4	32+40
	0917003	大数据导论	B	18	2	20+10
	0921107	Linux 操作系统	B	18	2	18+18
	合计学时					
第三学期	1200037	毛泽东思想和中国特色社会主义理论体系概论	A	18	2	28+8
	2100005	安全教育	A	√	√	4+0
	1400009	体育	B	18	2	4+24
	1200028	形势与政策	A	√	√	8+0
	0914019	MySQL 数据库	B	18	4	28+44
		Shell 编程	A	18	2	15+15
	0917006	Python 程序设计	B	18	4	30+42
	0917005	Hadoop 系统基础	B	18	6	44+64
合计学时						<b>198+226=424</b>
第四学期	1200029	形势与政策	A	√	√	8+0
	2100006	安全教育	A	√	√	4+0
	1200038	毛泽东思想和中国特色社会主义理论体系概论	A	18	2	28+8
	1200034	大学生职业发展与就业指导	A	18	1	16+5
		HBase 系统基础	B	18	4	32+40
		Hive 数据仓库系统	B	18	4	32+40
		Spark 大数据技术与应用	B	18	6	54+54
		大数据可视化技术与应用	B	18	2	18+18
	0911007	专业英语	B	18	2	30+0
合计学时						<b>222+165=387</b>
第五学期	2100007	安全教育	A	√	√	4+0
	0411164	跟岗实习	C	18	15w	0+390
	0411137	毕业设计（论文）	C	18	3w	0+78
	合计学时					
第六学期	2100008	安全教育	A	√	√	4+0
	0411142	顶岗实习	C	20	20w	0+520
	合计学时					
合计		实践学时数		1669	总学时	2554

	<b>实践学时所占比例</b>	<b>65.3%</b>
<b>说明:</b> 1.课程类型: A类(理论课) B类(理论+实践课) C类(实践课) 2.课程代码为教务管理系统中的课程代码,同一课程在不周学期开设使用不同代码。		

## 八、教学进程总体安排

表 8-1 教学进程安排表

课程结构	序号	课程名称	学时			考核方式	学时分配						学分		
			总学时	理论	实践		第一学年		第二学年		第三学年				
							第一学期 16周	第二学期 18周	第三学期 18周	第四学期 18周	第五学期 18周	第六学期 20周			
公共基础课程	1	国防教育与军事训练、入学教育	48		48	综合评价	2w								2.5
	2	思想道德修养与法律基础	54	40	14	过程考核+测试	2	2							3
	3	毛泽东思想和中国特色社会主义理论体系概论	72	56	16	过程考核+测试			2	2					4
	4	形势与政策	32	32		综合评价	√	√	√	√					2
	5	心理健康	14	14		综合评价	1								1
	6	安全教育	24	24		综合评价	√	√	√	√	√	√			2
	7	体育	84	12	72	过程考核+测试	2	2	2						4.5
	8	大学语文	60	60		过程考核+测试	2	2							3
	9	应用数学	84	72	12	过程考核+测试	4	2							4.5
	10	基础英语	84	84		过程考核+测试	2	4							4.5
	11	计算机应用基础	56	12	44	过程考核+测试	4								3
	12	大学生职业发展与就业指导	42	32	10	过程考核+测试	1			1					2.5
	13	创新创业教育	32	12	20	综合评价		2							2
<b>小计</b>			<b>686</b>	<b>450</b>	<b>236</b>		<b>18</b>	<b>14</b>	<b>4</b>	<b>3</b>				<b>38.5</b>	
专业课程	1	Java 程序设计(一)	56	28	28	技能测试	4								3
	2	HTML 5 网页设计	28	16	12	过程考核+测试	2								2
	3	Linux 操作系统	36	18	18	技能测试		2							2
	4	Java 程序设计(二)	72	32	40			4							4
	5	MySql 数据库	72	28	44	过程考核+测试			4						4
	6	Python 程序设计	72	30	42	过程考核+测试			4						4

	7	Hadoop 系统基础	108	44	64	过程考核+测试			6			6
	8	HBase 系统基础	72	32	40	过程考核+测试			4			4
	9	Hive 数据仓库系统	72	32	40	过程考核+测试			4			4
	10	Spark 大数据技术与应用	108	54	54	过程考核+测试			6			6
	11	大数据可视化技术	36	18	18	过程考核+测试			2			2
	12	毕业设计（论文）	78	0	78	过程考核+测试				3w		3
	13	跟岗实习	390	0	390	过程考核+测试				15w		15
	14	顶岗实习	520	0	520	过程考核+测试					20w	20
小计			1720	332	1388		6	6	14	16		79
专业拓展课程	1	计算机组装与维护	28	18	10	过程考核+测试	2					1.5
	2	计算机网络基础	30	20	10	过程考核+测试		2				1.5
	3	大数据导论	30	20	10	过程考核+测试		2				1.5
	4	Shell 编程	30	15	15	过程考核+测试			2			1.5
	5	专业英语	30	30	0					2		1.5
小计			148	103	45		2	4	2	2		7.5
选修课程	1	公共选修课 1										
	2	公共选修课 2										
	3	公共选修课 3										
	4	公共选修课 4										
小计			100									4
合计			2554	885	1669		26	24	20	21		129
<p>说明：</p> <p>1.校内外集中实训、毕业设计、顶岗实习周学时按 26 学时计算；</p> <p>2.标示“√”课程不占用正常教学时间，以讲座形式开展；</p> <p>3.公共选修课学时不计入总学时，只计学分。</p>												

表 8-2 可开设的非限定性专业选修课一览表

课程名称	课程类型	教学周数	建议周学时	学时数	学分
计算思维和数据科学	A	13	2	25	1
Java Web 应用开发	A	13	2	25	1
虚拟化技术与 docker	B	13	2	25	1
数据挖掘基础	B	13	2	25	1

表 8-3 集中实践教学项目一览表

实践教学项目	开设学期	开设地点	教学周数	总学时数

## 九、毕业标准

### （一）学分要求

学生须修完本专业培养方案中公共学习领域课（38.5 学分）、专业课程（79 学分）、专业拓展课程（7.5 学分）、公共选修课（4 学分），总学分达到 129 学分。

必修课、公共选修课（其中面授选修课 1 门，网络选修课三门或 60 课时）成绩合格。

### （二）素质要求

三年修业期间，素质拓展达到合格标准，取得学院颁发的素质评定证书。

### （三）职业资格证书要求

毕业前需取得以下职业资格证书或技能等级证书。

类别	资格证（技能证）名称	考核等级	考核学期	要求	职业编码
通用资格	大数据分析员	初级			
职业资格	大数据运维工程师	中级			
	大数据实施工程师				
	大数据采集工程师				
	腾讯云大数据认证				
	IBM 认证数据·工程师				
	阿里云大数据专业认证				

## 十、实施保障

### （一）师资队伍

本方案实施需要建立由专业带头人、骨干教师、“双师素质”教师、企业技术专家或技术能手共同组成的教学团队，生师比建议不高于 18:1；具有硕士学位教师占专任教师的比例达 35%以上；具有高级职务教师占专任教师的比例达 30%以上；专业基础课和专业课中双师素质教师比例达 70%以上；兼职教师承担实践课学时数比例达 50%。

#### 1. 专业带头人

校企各配置 1 名专业带头人。校内专业带头人应具有副高级以上技术职称，要有一定的云计算大数据专业知识，熟悉行业企业最新技术动态，能够站在云计算大数据专业领域发展前沿，把握专业技术改革方向，对本专业的前沿动态、行业发展、岗位需求等有较深入的了解，能够准确把握该专业建设与教学改革方向，具有对本专业发展的规划能力。

## 2. 骨干教师

专业教学团队应配置骨干教师 4 名以上。骨干教师应具有中级及以上职称，要能够根据行业企业岗位群的需要开发课程、编写教材，及时更新教学内容；承担 2 门以上专业课，具有课程开发及教学设计的能力，能够合理利用各种教学条件，采用不同教学方法和手段组织教学；能够开发校本教材、实训指导书，制作多媒体教学课件，建设精品网络资源共享课；到校企合作企业挂职锻炼，熟悉毕业生所从事工作岗位的要求。骨干教师必须是“双师素质”教师。

## 3. “双师素质”教师

“双师素质”教师应具有高等学校助理讲师（或以上）教师技术职务，年度考核合格，又具备下列条件之一：近五年有六个月（可累计）以上企业工作经历；近五年主持（或主要参与）2 项应用技术研究，成果已被企业使用，效益良好；近五年主持（或主要参与）两项校内实践教学设施建设或提升技术水平的设计安装工作，使用效果好，在省内同类院校中居先进水平；具有中级（或以上）工程系列专业技术职称或国家注册执业资格证书、职业资格证书者。其他情况可由学院教学指导委员会认定。

## 4. 兼职教师

企业兼职教师应具有熟练的大数据相关岗位技术能力和一定的教学水平，从事相关岗位工作 3 年以上；具有中级以上专业技术职务或高级工以上职业资格或在本行业享有较高声誉、具有丰富实践经验和特殊技能的“能工巧匠”；企业兼职教师上课或担任学生实践指导任务前，需经过教育教学培训；企业兼职教师承担专业实践课及顶岗实习学时数达 50% 以上，形成稳定的企业兼职骨干教师队伍。

## 5. 本专业教师知识、能力与素质要求

知识要求：

- (1) 掌握计算机的基本结构、工作原理，计算机网络的基本概念和技术；
- (2) 掌握软件工程的基本知识，熟悉软件工程每个阶段的任务和工具；
- (3) 掌握操作系统的特点及功能，熟悉存储系统、网络系统的结构和原理；

- (4) 掌握数据库系统的特点及功能，熟悉数据库表的设计和操作系统；
- (5) 熟悉面向对象的程序设计方法，包含 C#语言、JAVA 或 .NET 语言、HTML 语言、PHP 语言、Python 语言；
- (7) 熟悉分布式系统的特点和核心技术，熟悉 HTTP、REST、SOAP 协议；
- (8) 熟悉大数据相关技术，包括数据获取和预处理、数据存储、数据应用开发、数据挖掘和分析等；
- (9) 熟悉大数据相关系统的技术架构、工作原理和使用技术，包括 Hadoop、HBase、Hive、Spark、Mahout 等；
- (10) 熟悉虚拟化、云计算相关的技术；
- (11) 熟悉信息安全相关标准及法规。

能力要求：

- (1) 具有设计、开发、测试和部署 Web 应用的能力；
- (2) 具有为软件系统开发 REST 接口的能力；
- (3) 具有使用脚本搭建大数据技术相关系统的能力，包括 Hadoop、HBase、Hive、park、Mahout 等；
- (4) 具有使用相关工具对数据进行预处理的能力；
- (5) 具有设计、建立和使用数据仓库的能力。

素质要求：

- (1) 拥护党的领导，拥护社会主义，热爱祖国，热爱人民；热爱教育事业，具有良好的师德风范；
- (2) 掌握教育学理论，具备在教学中实施行动导向教学法的能力，灵活运用案例及项目教学法和任务驱动等方法实施课程教学；
- (3) 具有教学设计能力、课堂教学能力、指导实训项目的能力等较高的教学技能；
- (4) 具备一定的科研素养；
- (5) 具备提高自身专业素质的能力，适应大数据技术的快速发展；
- (6) 具有较强的敬业精神，具有强烈的职业光荣感、历史使命感和社会责任感，爱岗敬业，忠于职守，乐于奉献。

## (二) 教学设施

### 1. 校内实践教学条件

为保证人才培养方案的顺利实施，需建成与课程体系配套的校内实训基地和理实



一体化教室，为理实一体化教学提供有力保障。校内实训室应有设备及实训功能见下表。

实训室名称	主要设备	实训功能
大数据实训室	计算节点服务器	对学生端提供服务支持
	Training 实训模块	操作系统级的真机模拟训练
	Reporter 报表模块	大数据分析报表实训
	VmWeb 在线虚拟化桌面模块	虚拟机实训
	Dashboard 虚拟机控制台模块	虚拟机管理
	Compute 计算模块	虚拟机配置
	Networking 网络模块	配置内部拓扑，网络划分实训
	ShiroX 权限控制模块	用户身份验证和权限控制管理
	Swift 分布式存储模块	静态数据的分布式存储
	BlockStorage 块存储模块	数据库和文件系统
	Images 镜像管理模块	VM 镜像
	Hadoop 2.7.1	Hadoop 实训
	Python2.7.12	Python 相关的大数据实训
	Pandas 0.9.12	数据分析实训
计算机	学生操作用机实训	
网络管理实训室	服务器	文件共享
	二层交换机	交换机配置实训
	堆叠模块	交换机堆叠实验
	模块化路由器	路由器配置实训
	路由器线缆	路由器实验
	三层交换机	交换机配置实训
	投影仪	多媒体教学
	计算机	学生操作用机
	网络机柜	机房组网实验
	交换机	机房组网、交换实验
	信息安全攻防平台	网络安全实训
	防火墙及配件	防火墙实训
	路由器及配件	路由实验
网站开发实训室	服务器	ERP 平台服务器
	计算机	学生操作用机实训
	二层交换机	机房组网实训
	三层交换机	机房组网实训

实训室名称	主要设备	实训功能
	投影仪	多媒体教学
组装维护实训室	计算机	计算机组装、维护实训
	组装工具包	
应用软件开发实训室	服务器	文件共享
	计算机	学生操作用机
	电子白板	教师教学、演示
	打印机	教师教学、实训
	投影机	多媒体教学
	平板电脑（苹果及安卓）	学生项目实训、演示
	笔记本电脑	用于苹果 IOS 系统软件开发
基础实训室（6~8 个）	计算机	软件教学

## 2. 校外实习基地及要求

校外实训基地的主要功能：有利于学生掌握岗位技能、提高实践能力；满足学生半年以上顶岗实习的需要，从而实现学生在基地的顶岗后就业。校外实训基地有健全的规章制度及基于职业标准的员工日常行为规范，有利于学生在实训期间便养成遵纪守法的习惯，能真正地领悟到团队合作精神，同时能培养学生解决实际问题的能力。校外实习基地业有利于学校及时了解社会对人才培养的要求，及时发现问题，有针对性地开展教育教学改革。

本专业应通过政府、行业协会等平台或积极主动联系等方式，紧密联系行业企业，多渠道筹措资金，多形式开展合作。在校外实训基地建设中，积极寻求与国内外、区域内大型知名企业开展深层次、紧密型合作，建立 3~5 个相对稳定的校外实训基地，充分满足大数据技术与应用专业学生的综合实践能力及半年以上的顶岗实习需要。发挥企业在人才培养中的作用，由企业提供场地、办公设备、项目和技术指导人员，企业技术人员与教师共同组织和带领学生完成真实项目设计、施工、调试与维护，使学生真正进入企业项目实战，形成校企共建、共管的格局。

在学生顶岗实习过程中，每个实习基地聘用一名企业兼职联络员，负责所及提供企业生产信息，根据企业生产内容和专业教学进度协调实习实训相关事宜。专业应与各合作企业共同动态修订人才培养目标，共同制定实习方案，按照人才培养方案，聘请企业专家担任兼职教师，在专兼职教师的共同指导下，以实际工作项目为主要实习任务。学生通过在企业真实环境中的实践，积累工作经验，具备职业素质综合能力，

达到“准职业人”的标准，从而完成从学校到企业的过渡，使学生达到企业用人标准。

### （三）教学资源

#### 课程资源建设：

大数据技术与应用专业是一个新专业，教学资源很缺乏，为了保证教学质量，有必要开发建设一套专业教学资源库。资源库的建设应该从专业核心课程开始，逐步带动其他课程的教学资源建设，并在教学实践中不断完善。每一门课程应主要建设以下资源：

1. 课程标准，应说明课程定位、教学学时，对于前序课程的要求，以及后续课程的衔接，课程的教学目标、课程设计思路、课程内容及要求、教学手段方法及课程考核方式建议等内容；
2. 课程的教学课件，包含 PPT、例题程序代码；
3. 理实一体化课程教案，详细的教学步骤设计和教学内容；
4. 实验题目、实验内容、实验环境、实验说明书和程序代码；
5. 练习题及其参考答案；
6. 实训任务清单（实训环境、实训说明、实训要求、实训内容、实训考核方法）、实训项目代码、实训指导书；
7. 考试标准、试题库和参考答案。

#### 教材建设：

教材是教学内容的载体，可以呈现教学标准的内容，也可以体现教学方法。内容适度、结构合理的教材是教学质量保证的重要因素，建议从以下几方面加强教材建设。

##### 1. 校企合作共建“理实一体化”教材

专业组教师要联合企业一线技术专家，紧贴大数据当前实际，合作完成教材编写。

教材要将真实项目引入教材，实现理论知识学习和实际应用一体化；教材要面向教学过程、结合学生实际合理设置理论教学和技能训练环节，实现“教、学、做”甚至是“教、学、做、考”合一。

教材以项目为核心，每一教学单元建议采用教学导航、课堂知识讲解、课堂实践、课外拓展的环节开展教学。教学单元结束后，通过“单元实践”进一步提升技能；相关课程结束后，通过“综合实训”提升学生的综合能力。

##### 2. 选用高职高专优质或国家规划教材

### 网络资源建设：

以信息技术为手段，以网络为平台，构建体系完善、资源丰富开放式的专业教学资源。同时要善于整合、消化、吸收企业优秀教学资源，使其实行共享。

网络资源需从以下几方面进行建设：

#### （1）专业建设方案

专业建设方案包括：专业简介、专业人才培养方案、课程标准、教学文件等。

#### （2）职业技能标准

#### （3）精品资源共享课资源

①基本资源。基本资源应包含课程简介、课程标准、教学大纲、授课计划、教案、多媒体课件、学习指南、习题、实验实训项目、电子教材、试题库等。

②拓展资源。拓展资源是在基本资源基础上，面向学生和社会学习者扩展的自学、培训、进修、检索、科普、交流等内容，体现课程技术特点并向产业领域扩展。拓展资源包括素材库、培训包、工种包、企业案例、参考网站等。

③课程视频。课程视频包括课程整体设计介绍、课程单元设计说明等课程设计指导，课堂授课、现场教学、实训实习等教学场景，原理结构、工作过程、业务流程、操作步骤、技术细节、安全禁忌等内容。

#### （4）人文素养教学资源

①品德德育教学资源库。包含思想道德修养与法律基础、毛泽东思想和中国特色社会主义理论体系概论、职业生涯规划与创业就业指导等课程的课程资源，思政网站等。

②基础文化课教学资源库。包含本专业开设的高职语文、高职数学、高职英语、计算机应用基础、体育等文化基础课程的课程标准、教材、课件、案例库、习题库、视频资料等教学资源。

③职业拓展教学资源库。包含本专业开设的心理健康、形式与政策、拓展学习领域课程及公选课等课程课程资源。

### （四）教学方法

课堂教学的组织管理。系（部）聘任有相应学识水平、有责任心、有教学经验的专任或兼职教师任课。组织任课教师认真研究课程标准，组织编写或选用与标准相适应的教材和教学参考资料；要求教师认真履行教师岗位职责，按教学规律讲好每一节课；组织教师开展教学方法的讨论和研究，合理使用现代化教学手段，充分利用教学资源，保证课堂教学质量。理实一体化课程及实践性教学内容要严格按人才培养方案

和课程标准的要求进行教学，充分发挥校内外实训基地的教学资源，任课教师要设计好每一节或每个项目的教学做环节，训练学生的专业基本技能和综合职业能力。

对于公共基础课，建议采用启发式授课方式，一讲授为主，配合简单实验，多采用案例法、推理法等，深入浅出地讲解理论知识，注重学生参与，可制作图表和动画，易于学生理解。

对于基本技能课和岗位能力课，建议采用训练考核的教学方法，在讲清原理的基础上以实践技能培养为目标，保证训练强度达到训练标准，实践能力达到技术标准。可采用情景化教学、互动教学，教师可进行演示，为学生提供较为详尽的训练指导、动画视频等演示资料。

对于综合能力课，可采用项目教学法、案例教学法、任务驱动教学法、模块化教学法等，按照项目获任务的实施流程展开教学，让学生间接学习工程项目经验。项目教学法尽量配合小组教学法，可将学生分组教学，并在分组中分担不同的职能，培养学生的团队合作能力，积极指导学生学以致用，组织学生自主完成工作任务。

### （五）教学评价

凡是培养方案规定开设的课程都要对学生进行考核评价。根据课程特点和性质采用多样化的考核方式和方法，考核重点放在学生的综合素质及能力的评价方面，注重过程性评价。

课程考核与评价建议采用多方面结合的形式，应采用过程考核与结果考核相结合，理论与实践相结合，笔试、口试、操作相结合，离线作业、在线作业、终结性考核相结合，校内教师评价与企业教师评价相结合，他评、自评、互评相结合，根据课程不同特点，各考核项分配不同的比例进行成绩评定。

### （六）质量管理

#### （1）教学管理

①日常教学管理。为保证人才培养方案的有效实施，按照教务处统一的教学运行文件，教务处及系（部），对学院教学运行进行日常检查、抽查、和学期检查。一般采取听课、检查任课教师的教学文件、召开学生座谈会、对学生进行问卷调查等形式，对出现的问题及时纠正改进，以确保方案的正常运行。

②建立教学工作例会制度。根据学院教学工作需要，由教务处协助主管教学副院长定期和不定期召开教学工作会议，全体系（部）主任及相关部门人员参加。通过教学工作例会，传达并学习最新职教发展动态和教学改革理念，布置学院教学发展改革任务，了解系（部）日常教学及专业、课程建设工作进展情况，研究和处理人才培养

方案执行中出现的各种问题等。

③系（部）教学管理。系（部）定期召开专业主任会议和任课教师会议，及时掌握教学过程情况，总结教学工作和教学管理工作经验，及时研究解决教学过程中出现的问题。各专业要在每学期初制定出工作计划，组织集体备课、观摩教学、开展教学研究，了解教师教学进展情况，按学院要求进行教学检查。

## （2）教学质量监控体系

### ①教学督导委员会组织机构

建立院系两级教学督导委员会，分级管理，分工负责，协同监控。

院级教学督导委员会由学院党委书记任主任，分管教学工作和学生工作的两位副院长任副主任，同时聘请具有丰富教学经验的在职或离退休教师、具有丰富管理经验的的教学管理人员组成山西职业技术学院教学督导委员会。院级教学督导委员会由督导中心牵头，以教学目标和主要教学环节的宏观监控为主，在院领导的直接领导下，负责全校教学质量监控工作的总体协调，确保教学质量的稳步提高。主要工作职责：一是对专业设置的论证、专业人才培养方案及相关教学文件的审核；二是通过深入课堂、实验室、实习基地，客观掌握教学运行的全过程，提出督导建议，为学院有关教学决策提供参考依据。

系级教学督导委员会由系主任负责，成立由校企合作工作委员会和专家、优秀毕业生代表组成的人才培养质量监控小组。系级教学督导委员会的主要职责：以教学过程自我监控为主，在主要负责人的领导下，负责对本单位的整体教学工作、教师的教学情况、学生的学习情况进行监控。负责组织各专业的听课、试卷命题、阅卷、试卷质量分析、毕业论文质量分析等工作，并通过学院、系部、专业教研室组织的各类检查评估（教案、作业布置与批改、教学进度计划、学生评教、教师评学、教研活动的开展等），严把各个教学环节的质量。

### ②日常教学督导

听课制度：院级领导每月听课次数不少于 1 次；值班中层干部每周听课不少于 1 次；系（部）主任、副主任及系（部）书记每月听课不少于 2 次。学院和系（部）各级党政干部深入教学第一线，及时了解教学情况，倾听师生意见，发现并解决教学中存在的问题，避免教学一线与管理层的脱节，保证教学管理工作的针对性和有效性。

学生教学信息员制度：以专业班级为单位，确定思想品德优良，有参与教学管理的积极性，善于联系老师和同学，能客观反映广大学生的意见学生代表和学生干部，举行学期座谈会，填写任课教师评分表，给学生以畅通的渠道反映本系、本专业的教

学管理、办学条件和教学质量中存在的问题并对教学提出意见和建议，使系部的管理和教学更加贴近学生、贴近实际。

**教学检查与管理制度：**从学期初到学期末，院、系两级安排不少于 2 次的集中教学检查，采取听（听课、召开座谈会听取师生的反映）、看（查看教学条件和管理软件）、查（抽查教案、学生作业、实验报告、实习报告、课程设计、毕业设计等）、评（对教学条件、状态、效果进行评价）。教学情况的检查工作贯穿始终，发现问题及时反馈并解决落实。

为保证人才培养方案的执行，在教学运行中严格执行学院制定的教学工作规范、教学计划、课程标准和教学进程，严格教学事故的认定与处理，严格执行教学评价制度，严格执行课堂教学和实践教学过程的检查制度，严格教学文件的规范管理，保证人才培养方案的顺利实施、教学秩序的稳定和教学质量的提高。

## 附件 1 大数据技术与应用专业人才需求调研报告

### 一、调研目的与对象

#### （一）调研目的

通过本次调研收集和分析大数据专业学生的社会人才需求状况、能力要求以及相关岗位的工作内容，了解社会、行业以及企业对大数据技术与应用专业人才知识、技能、素质要求的变化和趋势，为我院大数据专业设置、教学教研、招生规模、学生就业指导提供信息，为专业人才培养目标定位、教学计划和课程标准的修订、教学的改革提供依据和帮助，提高学校人才培养质量及毕业生的就业质量。

#### （二）调研对象

此次调研共涉及山西多元合创教育科技有限公司、新华三集团、锐捷网络公司、北京西普阳光教育科技股份有限公司等多家企业。

调研组成员：陈炯，吴克强，苏彬，史志杰，王小洁，张剑，解莹，朱壮普，樊斌峰，王新萍，张爱民

调研时间：2019年5月

### 二、调研方法与内容

#### （一）调研方法

本次调研主要以问卷调查和专家座谈的方式进行，并结合网络企业招聘需求，调查了多家企业的大数据方向用人需求。

#### （二）调研内容

##### 1. 行业与就业调研

- （1）大数据国家及我省政策分析
- （2）大数据行业现状分析；
- （3）大数据行业前景分析；
- （4）大数据人才的需求分析。

##### 2. 企业调研

- （1）企业招聘员工的素质要求；
- （2）企业大数据从业人员的主要岗位；
- （3）企业大数据岗位的主要工作任务；
- （4）企业大数据岗位的职业发展路径。



### 三、调研分析

#### (一) 行业调研

##### 1. 国家及我省政策

2016年，习近平总书记指出，“大数据战略作为‘十三五’期间的十四大国家战略之一，是中国经济发展新的驱动力。”至此，大数据战略上升到国家战略。未来五年信息化建设将重点实施网络强国战略，并且明确提出实施国家大数据战略。

工业和信息化部为全面部署“十三五”时期大数据产业发展工作，加快建设数据强国，为实现制造强国和网络强国提供强大的产业支撑，历时一年多组织编写了《大数据产业发展规划（2016-2020年）》。通过定量和定性相结合的方式提出了2020年大数据产业发展目标。在总体目标方面，提出到2020年，技术先进、应用繁荣、保障有力的大数据产业体系基本形成，大数据相关产品和服务业务收入突破1万亿元，年均复合增长率保持30%左右。在此基础之上，明确了2020年的细化发展目标，即技术产品先进可控、应用能力显著增强、生态体系繁荣发展、支撑能力不断增强、数据安全保障有力

国务院印发《促进大数据发展行动纲要》通过促进大数据发展，释放技术红利、制度红利和创新红利，倒逼行政体制改革，提升政府治理能力，推动经济转型升级。

新一届山西省委、省政府把大数据战略摆在了非常重要的位置。山西作为资源大省，是国务院设立的“国家资源型经济转型综合配套改革试验区”，时逢加快经济结构调整、实现转型跨越发展的关键时期，实施大数据战略更具特殊意义。2016年召开的山西省第十一次党代会，将“实施大数据战略、发展数字经济、建设智慧山西”提上了日程。目前，山西省已成立了由省长楼阳生任组长的大数据发展领导小组，我省目前已出台《山西省促进大数据发展应用的若干政策》、《山西省大数据发展规划（2017-2020年）》、《山西省促进大数据发展应用2017年行动计划》等多项扶持政策，通过优先安排项目用地供应、给予企业0.35元/千瓦时优惠电价、减免税收红利政策、设立50亿元规模发展引导基金吸引大数据企业落户山西。目前，华为、中兴、浪潮、新华三以及三大运营商已经或即将与我省签署战略合作协议，吕梁国家数据产业园、华得云数据中心、中交网通西部数据中心等30个项目已经在我省落地或者达成了投资意向。山西“大数据”时代已经全面来临。

(1) 2012年8月百度云计算（阳泉）中心奠基动工，目前，部分项目已投入运行，服务器运行数量为8万台。

(2) 2014年6月份吕梁云计算中心，国防科技大学和吕梁市政府重点合作建设项目正式运行，目前已为部分省直单位和吕梁市提供电子政务信息化服务，并为全国160多家科研院所提供高性能计算服务。

(3) 2015年3月，《国务院关于积极推进“互联网+”行动的指导意见》(国发〔2015〕40号)正式印发。

(4) 2015年6月，工信部提出要加快推进云计算与大数据标准体系建设(第七届云计算大会)

(5) 2015年7月，《国务院办公厅关于运用大数据加强对市场主题服务和监管的若干意见》(国办发〔2015〕51号)正式发布。

(6) 2015年8月，《促进大数据发展行动纲要》(国发〔2015〕50号)正式发布。

(7) 2016年7月，山西焦煤集团公司与山西联通达成“互联网+智慧能源”战略合作，共同建设山西“能源云”数据中心，探索互联网与能源行业融合，推进山西省能源信息化发展水平。

(8) 2016年7月，华为公司与吕梁市政府签订了云服务大数据中心合作协议。项目预计投资7个亿，计划于2017年投入运营，立足于打造山西省云计算产业支撑中心。

(9) 2016年12月，《大数据产业发展规划(2016-2020年)》(工信部规〔2016〕412号)正式由工业和信息化部发布。

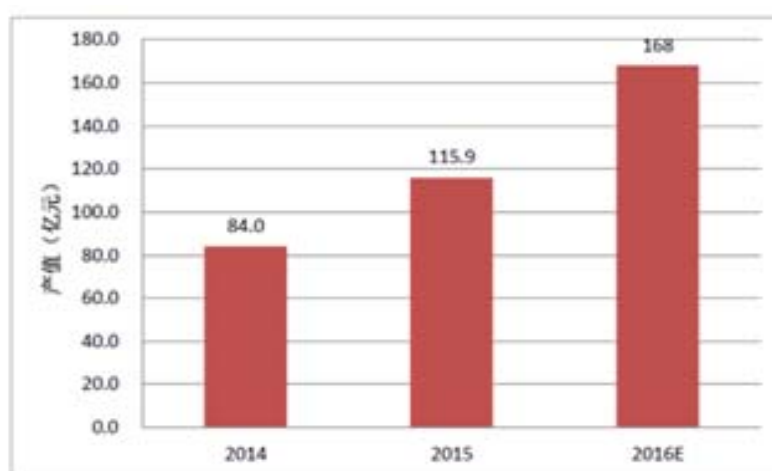
(10) 2016年12月，中国信息通信研究院举办了“2017年ICT深度观察大型报告会暨白皮书发布会”，会上正式发布了《大数据白皮书(2016)》。

## 2. 行业现状

大数据是新资源、新技术和新理念的混合体。从资源视角来看，大数据是新资源，体现了一种全新的资源观。1990年以来，在摩尔定律的推动下，计算存储和传输数据的能力在以指数速度增长，每GB存储器的价格每年下降40%。2000年以来，以Hadoop为代表的分布式存储和计算技术迅猛发展，极大的提升了互联网企业数据管理能力，互联网企业对“数据废气”(Data Exhaust)的挖掘利用大获成功，引发全社会开始重新审视“数据”的价值，开始把数据当作一种独特的战略资源对待。大数据的所谓5V特征(数据量大、结构多样、数据价值密度较低、增长速度快、数据真实性较高)主要是从这个角度描述的。

目前大数据产业的统计口径尚未建立。对于我国大数据产业的规模，各个研究机构

均采取间接方法估算。中国信息通信研究院结合对大数据相关企业的调研测算，2015 年我国大数据核心产业的市场规模达到 115.9 亿元，增速达 38%，预计 2016 年将达到 168 亿元，2017-2018 年还将维持 40%左右的高速增长。



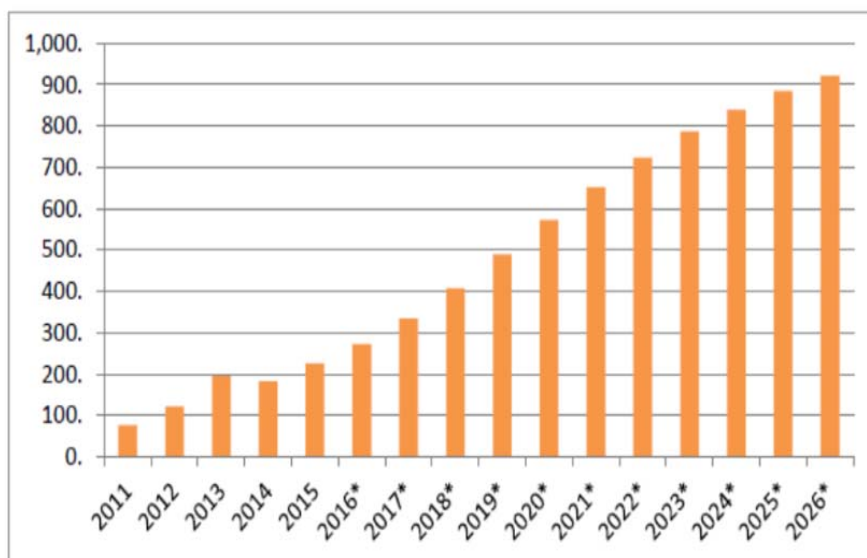
数据来源：中国信息通信研究院，2016 年 8 月，单位：亿人民币

图 1 中国大数据产业规模

我国已经初步形成了由互联网企业（以百度、阿里、腾讯为代表）、传统 IT 厂商（以联想、华为、浪潮、曙光、用友等为代表）、大数据企业（深圳视界信息技术有限公司（八爪鱼）、成都四方伟业软件股份有限公司、智慧星光信息技术有限公司等为代表）共同组成的市场供给关系，但各环节发展水平不均衡，在大数据产业链高端环节缺少成熟的产品和服务，面向海量数据的存储和计算服务较多，而前端环节数据采集和预处理，后端环节数据挖掘分析和可视化，及大数据整体解决方案等产品和服务较为匮乏。

### 3. 行业前景

大数据是伴随着信息数据爆炸式增长和网络计算技术迅速发展而兴起的一个新型概念。在现实世界里，物质、能量、信息是共同构成现实在互联网及智能硬件快速普及的当下，数据以爆炸方式增长。据 Forrester 的研究结果，目前在线或移动金融交易、社交媒体、GPS 坐标等数据源每天要产生超过 2.5EB（1EB 为  $2^{60}$  次方）的海量数据。据 IDC 预计，全球数据总量增长率将维持 50%左右，2020 年全球数据总量将达到 40ZB；中国将达到 8.6ZB，占全球的 21%。根据中国信息产业研究院的数据显示，2015 年中国大数据市场规模约为 116 亿元，同比增长 38%；预计未来几年，随着应用效果的逐步显现，中国大数据市场规模还将维持 40%左右的高增长。根据 IDC 的最新预测，到 2019 年，来自大数据和业务分析的年度全球收入将从 2015 年的 1220 亿美元增至 1870 亿美元。



数据来源: Wikibon, 2016 年 3 月, 单位: 亿美元

图 2 全球大数据产业规模

数据的膨胀需要更多的数据中心来存储和处理数据，同时云计算、互联网的高速发展也需要硬件设施来提供支持。互联网+已经成为各个行业的发展趋势，IDC 业务必将成为云时代的“商业地产”，为互联网发展提供成长的土壤，在此领域耕耘的公司将最先并持续受益，同时也面临着相当激烈的市场竞争。国内 IDC 近年来发展迅速，未来几年增速将保持在 40%左右。

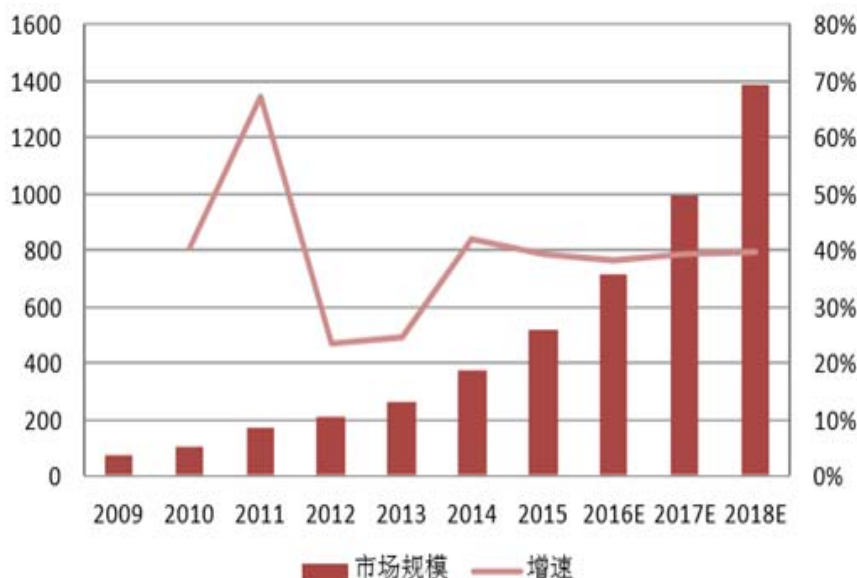


图 3 国内 IDC 市场规模及预测 (单位: 亿元)

日前，市场研究机构易观智库发布了《中国城市大数据市场专题分析》报告，报告称 2016 年我国城市大数据市场规模达 132.8 亿元，同比增长 45.9%，到 2017 年有望增

至 189.4 亿元。

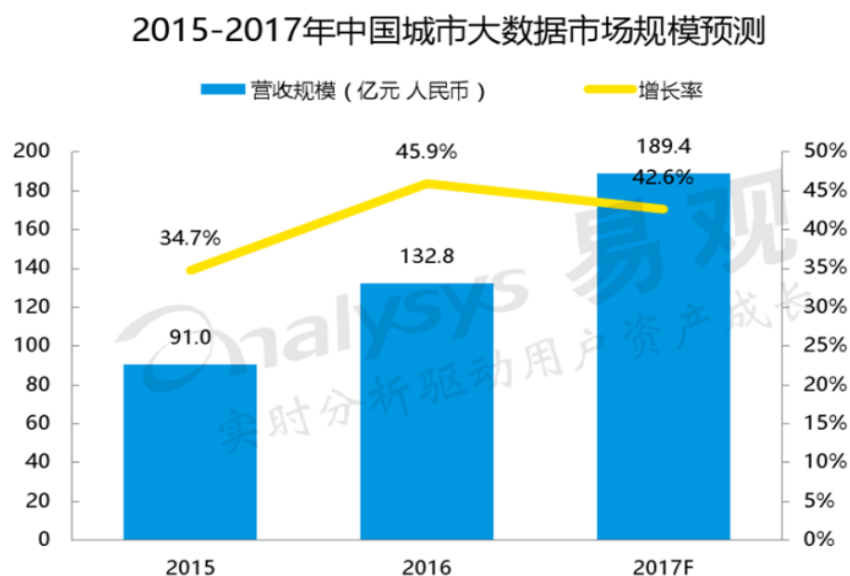


图 4 2015—2017 年中国城市大数据市场规模预测

报告同时指出，从市场细分领域来看，大数据软件和服务比重呈上升趋势，硬件比重逐年减少。与 2015 年相比，2016 年大数据硬件市场规模为 53.9 亿元，占比下降 1.8%；软件市场规模为 72.6 亿元，占比提高 0.8%；服务市场规模达 41.5 亿元，占比提高 1.0%。

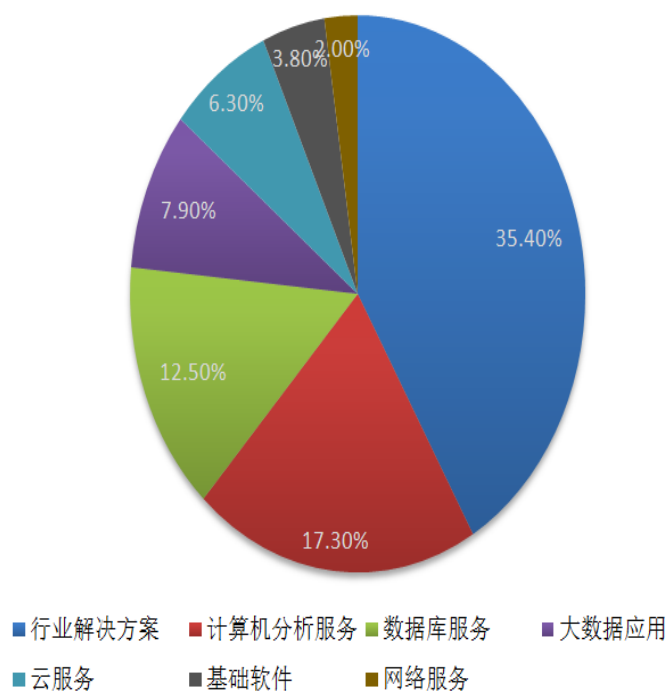


图 5 大数据市场细分比例

麦可斯—中国 2017 届大学毕业生培养质量跟踪评价报告从月收入、就业满意度、工

作与专业相关度，以及专业所对应的热门职业的就业质量差异等方面，细致分析 2017 届高职毕业生就业数量（就业比例）较大的十个专业的就业前景。

2017 届毕业生就业数量较大的十个高职高专专业分别是会计、工程造价、护理、机电一体化技术、建筑工程技术、会计电算化、物流管理、汽车检测与维修技术、计算机应用技术、电子商务。

2017 届高职高专十大专业的就业率、月收入、工作与专业相关度、就业满意度与离职率进行对比分析发现，从月收入方面来看，计算机应用技术专业月收入最高（4150 元）。

2018 高职类 10 大专业中，计算机应用技术专业和电子商务专业的就业满意度最高，均为 67%。

**表4 全国2017届毕业生就业数量较大的前十位高职高专专业分别对应的就业数量较大的前三位职业的就业情况**

TOP	专业名称	该专业就业数量较大的前三位职业	毕业半年后月收入(元)	工作与专业相关度(%)	就业满意度(%)
1	会计	会计	3355	97	65
		出纳员	3315	95	62
		文员	3129	20	55
2	工程造价	预算员	3334	98	66
		建筑技术员	3804	92	63
		施工技术员	3735	87	59

3	护理	护士	3547	99	65
		护士助理和护理员	3488	96	64
		紧急医疗救护及护理人员	3505	99	67
4	机电一体化技术	电气技术员	4146	90	65
		电气工程技术员	4237	90	63
		电气和电子运输设备安装者和修理技术员	4112	91	58
5	建筑工程技术	施工技术员	3793	96	64
		建筑技术员	3927	97	63
		土木建筑工程技术员	3810	96	61
6	会计电算化	会计	3201	96	69
		出纳员	3102	93	65
		文员	3055	20	60
7	物流管理	文员	3307	31	61
		存货管理员(储藏室、库房的)	3813	92	63
		物流专员	3797	95	60
8	汽车检测与维修技术	车身修理技术员	3024	95	66
		汽车机械技术员	3188	95	62
		其他销售代表、服务商	4294	64	77

9	计算机 应用技术	计算机程序员	5168	99	74
		互联网开发师	5579	94	75
		平面设计	3738	86	69
10	电子商务	电子商务专员	4181	94	73
		文员	3383	25	60
		客服专员	3821	66	61

表3 全国2017届毕业生就业数量较大的  
前十位高职高专专业的就业情况



TOP	专业名称	毕业半年后就业率 (%)	毕业半年后月收入 (元)	工作与专业相关度 (%)	就业满意度 (%)	毕业半年内离职率 (%)
1	会计	92.5	3476	62	64	46
2	工程造价	93.3	3632	68	63	41
3	护理	90.2	3374	92	64	17
4	机电一体化技术	92.7	4104	50	64	44
5	建筑工程技术	92.0	3393	71	62	42
6	会计电算化	92.6	3340	58	66	47
7	物流管理	92.1	3883	43	64	54
8	汽车检测与维修技术	92.0	3784	57	64	54
9	计算机应用技术	90.3	4150	53	67	50
10	电子商务	92.2	4028	51	67	55

## 四、结论与建议

### (一) 结论

根据当前我国尤其是我省大数据行业的发展现状和趋势，以及我院相关专业的人才培养格局，通过调研、分析与研究，得出的结论是，应当充分利用我院在信息技术、大数据研究方向的积累和优势，应加大大数据专业人才培养数量。

### (二) 建议

#### 1. 按照就业方向制定教学计划

按照高等职业教育的培养目标，转变教学理念，结合社会生产实际，明确大数据技术与应用专业的就业方向，制定适应高职教学要求的教学计划。教学计划应以提高人才质量为目标，以模式改革创新为主题，以培养学生的技术应用能力和职业基本能力为主线，以技能训练为重点，保证学生有上岗就业的职业资格。

#### 2. 按照职业能力进行课程设置

建立以工作过程为导向、能力培养为本位、素质教育和技能训练并重的课程体系，整合专业课程，加强技能训练，充分保证学生有良好的职业能力。

### 3. 按照岗位技能强化实践技能训练，改革教学模式

理论教学要按照岗位技能的知识需求进行，形成以项目或任务为导向，理实一体化、做中学的教学模式，要强化实验实训基地的建设，积极开展校企合作、工学结合、顶岗实习的工作，加大课程和实践场所的匹配，为课程改革提供基础保障，通过校企合作让企业文化尽早进入校园，让学生熟悉企业，让企业了解学生。

### 4. 对教师进行以提高技能水平为目的培训

教师实践技能的提高是提高高等职业教学水平的关键，要采取多种方法提高教师的实践能力和技能水平，鼓励教师下厂锻炼、学习企业最新技术，积极参与课题研究。

### 5. 注重职业素养和职业道德的培养

在专业教学中要贯穿职业素养和职业道德的培养，职业素养和职业道德的培养不是靠简单的说教就能解决的，一定要在专业教学中将这些要素融入进去，在教学中注重培养学生的质量意识、合作意识和严谨的工作态度，要求学生认真对待每一个项目，依次提高自身职业素养、质量意识、效益意识、团队合作精神、吃苦耐劳的精神等。

## 附件 2 职业标准

### 大数据运维工程师：

岗位职责：

1. 负责数据平台的运维管理工作（部署、监控、优化、故障处理）；
2. 负责 Hadoop/Spark/Flink/Elasticsearch/Kafka 等系统的架构审核、容量规划、

成本优化；

3. 负责大数据平台的用户管理、权限分配、资源分配；
4. 参与数据挖掘、机器学习平台的设计，并给出可执行的运维方案；
5. 参与数据平台的相关工具开发（包括自动化部署、监控、ETL 等）；
6. 深入理解数据平台架构，发现并解决故障隐患及性能瓶颈；
7. ETL 工具、调度工具、关系型数据库的运维。

任职资格：

1. 1 年以上大数据相关组件运维经验 (hadoop/yarn/hbase/hive/spark/kafka 等)，1 年以上的 CDH 或 HDP 维护经验，3 年以上系统运维相关经验；
2. 深入理解 Linux 系统，能独立部署开源软件，熟练掌握一门以上脚本语言 (shell/perl/python 等)，熟悉 python 开发语言优先；
3. 逻辑思维能力强，做事有条理，责任心强，工作积极主动，执行力强，有良好的团队协作意识。

### 大数据开发工程师

岗位职责：

1. 参与大数据分布式应用系统服务器端或客户端软件开发工作（需求开发、故障解决和性能优化等）；
2. 从事大数据技术研究和开发工作，跟进大数据技术发展方向；
3. 基于 hadoop 生态的大数据存储平台搭建和部署；
4. 承担数据抽取、清洗、转化等数据处理程序开发；
5. 商用平台指标数据挖掘。

任职要求：

- 1、计算机或者相关专业本科及以上学历，2 年以上相关工作经验；

- 2、CET-4，能熟练用英语阅读技术文档；
- 3、精通 java 语言开发，同时熟悉 groovy/scala 语言开发优先考虑；
- 4、熟悉 Hadoop 工作流程，熟悉 MR 执行流程，能够编写 MapReduce 程序，同时具有 storm/spark 开发经验优先考虑；
- 5、熟练使用 SQL，熟悉数据库原理，熟练使用至少一种主流关系型数据库；
- 6、熟悉 ETL 开发，能熟练使用至少一种 ETL (talend, kettle 等) 转化开源工具优先考虑；
- 7、熟悉 Linux 操作系统，熟练使用常用命令，熟练使用 shell 脚本；

## 附件 3 大数据技术与应用专业核心课程标准

## 《Java 程序设计(一)》课程标准

## 一、课程基本信息

课程名称	Java 程序设计(一)				
课程代码		学时	56	学分	4
授课时间	第一学期	适用专业	大数据技术与应用		
课程性质	岗位能力课程				
先修课程		后续课程	Java 程序设计(二)		

## 二、课程定位

本课程是计算机应用技术专业的岗位能力课程，目的是通过课堂教学和实践教学相结合，使学生能够深入理解面向对象概念，清楚的了解 Java 软件开发工作流程，建立起应用程序的概念，最终能够掌握 Java 软件开发的基本方法、基本技能，培养利用 Myeclipse、IntelliJ IDEA 等常用工具软件进行 Java 应用软件产品的分析、设计、编码、测试的综合应用能力，培养科学的思维方法，灵活运用知识的能力，实验操作能力，使学生具有较强的发现问题、分析问题、解决问题的能力，具有毕业后的直接上岗能力（或经短期培训后上岗），并且毕业时已具备一定的 Java 软件开发经验。

## 三、课程设计思路

本课程主要以教材为参考，融入经典的项目案例，同时借鉴了一些企业级项目模块的雏形，有利于学生的理解与学习。不失传统性，任何一门语言课程都要了解该语言的书写风格与运行的平台特点，但是在学习过程中，又主要通过项目实例让学生进一步学习基础知识同时也见到了项目雏形。通过前期近一百多课时的学习，总体上掌握该门课程的特点与学习思路，然后利用较综合性项目培养学生对软件组织结构、思想、过程的设计。新课程理念要求充分的调动学生思考和动手能力，先学后教，精讲多练。

## 四、课程目标

## (一) 能力目标

## (一) 能力目标

1. 培养学生规范编码和良好的程序设计风格；
2. 培养学生面向对象编程的思维和提高逻辑思维能力；
3. 培养学生发现问题、分析问题和解决问题的能力。

## (二) 知识目标

1. 初识 Java;
2. 掌握变量和数据类型;
3. 熟练掌握选择结构;
4. 熟练掌握循环结构;
5. 了解多重循环及程序调试;
6. 掌握一位数组的应用;
7. 了解二维数组的应用;
8. 吃货联盟项目;
9. 理解 Java GUI 概念;
10. 初识 Java Swing;
11. Swing 容器;
12. 掌握 Java Swing JList;
13. 掌握 Java Swing GridBagLayout;
14. 简单的登录注册项目;

### (三) 素质目标

1. 培养按时完成项目的良好习惯，增强集体和社会责任感;
2. 培养团队协作和良好的沟通能力;
3. 培养学生务实、创新、诚实、守信的性格。

### 五、课程内容及要求

序号	教学内容	能力目标	知识目标	教学方法及手段	学时
1	JAVA 开发环境的搭建和 MyEclipse 的安装与使用。	能够搭建开发环境，掌握 MyEclipse 的使用方法	1. 了解 Java 的发展历史 2. 了解 Java 程序的运行方式	引导启发学生独立完成设置工作，多媒体、上机操作、用 MyEclipse 工具做演示讲解一些基本概念	3
2	变量的声明和定义规范，数据类型的分类和应用，运算符的优先级和结合性	能够熟练定义和使用变量。根据数据类型的不同选择合适的应用场合。	1. 掌握 Java 语言的基本组成 2. 掌握数据类型、变量、运算符及表达式	由教师示范、指导，学生小组协作	3
3	流程控制：if 选择结构、多重 if 选择结构、嵌套 if 选择结构	掌握不同选择结构的代码结构和应用场合	1. 理解代码执行顺序 2. 掌握流程控制执行原理	由教师示范、指导、学生小组协作	4
4	循环结构：for 循环，while 循环，do-while 循环	掌握不同循环的代码结构和应用场合	1. 理解代码执行顺序 2. 掌握各种循环执行原理	由教师示范、指导、学生小组协作	4

5	双重循环和 Debug 程序调试	掌握多重循环的代码结构和应用场合。熟练使用 Debug 调试工具	1. 掌握多重循环执行原理 2. 掌握 Debug 调试工具的使用	由教师示范、指导、学生小组协作	5
6	一维数组	掌握一维数组的多种写法结构和遍历数组	1. 理解一维数组的储存原理 2. 了解一维数组遍历的运行机制	由教师示范、指导、学生小组协作	4
7	二维数组	掌握二维数组的写法结构和遍历数组	1. 理解二维数组的储存原理 2. 了解二维数组遍历的运行机制	由教师示范、指导、学生自主学习	4
8	吃货联盟	提升对技能点的运用能力, 积累项目经验。	1. 掌握 java 的基础知识, 理解项目的基本流程	学生自主学习探讨	8
9	GUI	了解 GUI 的概念, 能简述出来	1. 了解 java 与 GUI 的关系 2. 用到哪些包, 作用是什么	由教师讲解、学生自主学习	3
10	Swing	掌握 swing 的基本知识 理解 swing 下的组件的含义	1. 理解 swing 的作用 2. 能编写 swing 的简单代码	由教师示范、指导、学生自主学习	3
11	Swing 容器	理解容器的概念 知道需要考虑的一些注意事项	1. 知道常用的容器及作用	由教师示范、指导、学生自主学习	3
12	Java Swing JList	理解 JList 是什么 了解 JList 类的常用的方法	1. 理解 JList 的基础知识 2. 了解 JList 的几种配置	由教师示范、指导、学生自主学习	3
13	GridBagLayout	理解他是什么 知道他的作用	1. 学会如何简单使用 GridBagLayout	由教师示范、指导、学生自主学习	3
14	登录注册	提升对技能点的运用能力, 积累项目经验。	1. 根据 GUI 实现简单的登录注册 2. 可以对 GUI 更好的使用	学生自主学习探讨	8

## 六、课程实施建议

(一) 教学建议 (从教学条件、教学方法与手段、课程资源的开发与利用、教材选用等方面进行说明)

### 1. 教学资料开发建议

- (1) 结合实际软件开发工程案例, 进行课件制作、教学环境模拟。
- (2) 与企业专家合作编写符合技能培养的实验实训教材, 是学生学有所用。
- (3) 提供网络资源学习平台, 使学生不拘泥于固定的时间、地点学习, 提高学生

学习的灵活性。

## 2. 教学资源使用建议

(1) 学习包和教材，都是素材性的课程资源，但教材是知识的载体，而学习包是引导学生学习的载体。因此，要开发学习包，突出以学生为中心的学习过程，将本课程的教学内容与地区经济的发展密切结合，并体现学生职业生涯发展的需要。

(2) 教学中学生的问题、困惑、见解、情感和体验等都是动态生成的课程资源，老师应重视这些教学过程中动态生成的课程资源，并主动性和创造性地运用，使以项目为主的教学充满生机与活力。

(3) 本课程的实施，不仅应充分开发校内的课程资源，还应利用校外的课程资源，如相关企业的岗位实习和实践，尽量做到校企结合、工学结合。

(4) 积极开发教学资源：课程标准、实训指导书、授课计划等教学文件，以及课件、习题、案例库、网络方案、工具软件和网络资源等。

充分利用网络资源，了解最新的技术策略和实施方式，培养学生自主学习的能力。

## 3. 教材

选用教材：《Java 语言程序设计》 王新萍 电子工业出版社

### (二) 考核建议

考核主要由平时成绩、理论考试以及实训项目成绩组成。

#### 1. 平时成绩

平时成绩强调平时的出勤、课堂纪律、课堂表现和作业，教师通过千分制的实施可以在整个授课过程中监控学生的学习效果以及学习态度等等。平时成绩占全部考试的 30%。

#### 2. 理论考试

理论考试采用传统的笔试方式，为了更好的达到考核的目的，我们采用第三方出题的形式来组织理论考试。理论考试占全部考试的 40%。

#### 3. 项目评审

实施项目内部考核及教师考核相结合的原则，在项目实训末期，要进行项目评审，根据项目的完成情况给学生打分。项目评审成绩占全部成绩的 30%（包括学生自我评价及对同伴的评价、完成项目效果评价以及教师评价，其中自我评价(占 30%)、项目效果评价(占 30%)、教师评价(占 40%)。)

## 七、需要说明的其他问题（参考资料、所需仪器、设备、教学软件等）

### 1. 参考资料



- (1) 《Java 程序设计项目教程》 成奋华高等教育出版社
  - (2) 《Java 程序设计案例教程》钱银中编著机械工业出版社(国家级精品课程配套教材 21 世纪高职高专系列教材)
  - (3) 《Java 程序设计实例教程》马世霞机械工业出版社
  - (4) 与 Java 程序设计相关的其它书籍、网络资源
2. 硬件要求：微机、网络。
  3. 软件要求：JDK、SQL Server、idea、MyEclipse 等软件。

## 《Java 程序设计（二）》课程标准

## 一、课程基本信息

课程名称	Java 程序设计（二）				
课程代码		学时	72	学分	4
授课时间	第二学期	适用专业	大数据技术与应用		
课程性质	岗位能力课程				
先修课程	Java 程序设计（一）	后续课程	Mysql 数据库		

## 二、课程定位

本课程是计算机应用技术专业的岗位能力课程，目的是通过课堂教学和实践教学相结合，使学生能够深入理解面向对象概念，清楚的了解 Java 软件开发工作流程，建立起应用程序的概念，最终能够掌握 Java 软件开发的基本方法、基本技能，培养利用 Myeclipse、IntelliJ IDEA 等常用工具软件进行 Java 应用软件产品的分析、设计、编码、测试的综合应用能力，培养科学的思维方法，灵活运用知识的能力，实验操作能力，使学生具有较强的发现问题、分析问题、解决问题的能力，具有毕业后的直接上岗能力（或经短期培训后上岗），并且毕业时已具备一定的 Java 软件开发经验。

## 三、课程设计思路

本课程主要以教材为参考，融入经典的项目案例，同时借鉴了一些企业级项目模块的雏形，有利于学生的理解与学习。不失传统性，任何一门语言课程都要了解该语言的书写风格与运行的平台特点，但是在学习过程中，又主要通过项目实例让学生进一步学习基础知识同时也见到了项目雏形。通过前期近一百多课时的学习，总体上掌握该门课程的特点与学习思路，然后利用较综合性项目培养学生对软件组织结构、思想、过程的设计。新课程理念要求充分的调动学生思考和动手能力，先学后教，精讲多练。

## 四、课程目标

## （一）能力目标

## （一）能力目标

1. 培养学生规范编码和良好的程序设计风格；
2. 培养学生面向对象编程的思维和提高逻辑思维能力；
3. 培养学生发现问题、分析问题和解决问题的能力。

## （二）知识目标

1. 认识类与对象；

2. 掌握方法与方法重载；
3. 掌握封装与继承；
4. 掌握方法重写与多态；
5. 熟练掌握抽象类和接口；
6. 熟练掌握 switch 选择结构；
7. 掌握 Java 的异常处理；
8. 掌握集合框架和泛型（一）；
9. 掌握集合框架和泛型（二）；
10. 熟练掌握实用类（一）；
11. 熟练掌握实用类（二）；
12. 掌握 Java 的输入/输出（一）；
13. 掌握 Java 的输入/输出（二）；
14. 熟练掌握多线程；
15. 了解网络编程；
16. 熟练掌握 XML（一）；
17. 熟练掌握 XML（二）；
18. 嗖嗖移动业务大厅项目；

### （三）素质目标

1. 培养按时完成项目的良好习惯，增强集体和社会责任感；
2. 培养团队协作和良好的沟通能力；
3. 培养学生务实、创新、诚实、守信的性格。

## 五、课程内容及要求

序号	教学内容	能力目标	知识目标	教学方法及手段	学时
1	类的设计与对象的创建及使用：矩形类的设计与实现	掌握简单的类与对象设计与实现方法	1. 理解面向对象的程序设计方法 2. 了解类与对象	由教师示范、指导、学生小组协作完成矩形类的设计与实现	4
2	方法的使用和方法重载	掌握方法和发发重载的写法和区别及共同点	1. 掌握方法的定义 2. 掌握方法重载的定义	由教师示范、指导，学生小组协作	4

3	类的封装与继承：点、圆、圆柱类的设计与实现	掌握简单的类的封装与继承的概念设计与实现方法	理解类的封装与继承的概念及实现	由教师示范、指导、学生独立编程实现点、圆、圆柱类的设计与实现	4
4	类的方法重写与多态性：点、圆、圆柱类的设计与实现	掌握简单的类的方法重写与多态性的概念设计与实现方法	理解类的方法重写与多态性的概念及实现	由教师示范、指导、学生独立编程实现点、圆、圆柱类的设计与实现	4
5	抽象类的定义与接口的实现	掌握抽象类的定义规则和接口的定义与实现	1. 理解抽象类的概念和定义 2. 掌握简单接口设计与实现	由教师示范、指导、学生独立编程实现图形类的设计与实现	4
6	switch 的使用与应用场景	掌握 switch 的使用与应用场景	1. 掌握 switch 的使用。 2. 理解 switch 的应用场景	由教师示范、指导、学生独立编程实现图形类的设计与实现	4
7	异常的处理：标准异常与用户自定义异常	1. 掌握常见异常的处理 2. 学会用户自定义异常的创建	1. 理解异常的概念 2. 了解异常类的层次结构	由教师示范、指导、学生小组协作完成图形类的设计与实现	4
8	集合框架：List 集合和 Set 集合	掌握 List 集合和 Set 集合的概念和定义	1. 掌握 List 集合的特点和应用 2. 掌握 Set 集合的特点和应用	由教师示范、指导、学生自主学习	4
9	集合框架和泛型：Map 集合和泛型的应用	1、掌握 Map 集合的概念和定义 2、掌握泛型的使用规则	1. 掌握 Map 集合的特点和应用 2. 理解泛型的使用场景	由教师示范、指导、学生自主学习集合的定义与使用	4
10	String 与 StringBuffer 类：字符串的创建、比较、查找、转换、插入、删除等	掌握类的比较、查找、转换、插入、删除等方法	1. 理解 String 与 StringBuffer 类的概念、区别 2. 掌握二个类的常用构造方法	由教师示范、指导、学生自主学习	4
11	枚举和 Math 类、Date 类的多种方法	掌握枚举的定义和 Math 类、Date 类的多种方法	1. 掌握枚举的定义和应用场景 2. 掌握 Math 类、Date 类的应用	由教师示范、指导、学生小组学习	4
12	输入与输出：基本输入与输出	理解常用的输入输出流类及方法	理解输入输出流的基本概念	由教师示范、指导、学生自主学习	4
13	输入与输出：文件输入与输出	会用文件输入输出流类的基本方法设计程序	1. 理解文件输入输出的基本概念 2. 理解常用的文件输入输出流类及方法	由教师示范、指导学生自主学习	4
14	多线程的处理：用二种不同的方法来实现多线程	1. 理解线程的状态及生命周期，线程类及线程对象的创建 2. 会用输入输出流类的基本方法设计程序	1. 了解程序、进程、线程及多线程的概念 2. 基本掌握多线程程序的设计方法	由教师示范、指导学生自主学习	4

15	网络编程	1. 理解 C/S 架构。 2. Socket 编程思想以及应用	1. 了解 C/S 架构思想, C/S 架构与 B/S 架构的区别 2. 会编写客户端与服务端, 并实现客户端与服务端之间通信	由教师示范、指导学生自主学习	4
16	XML (一): DOM 对 XML 文档进行增删改查的操作	1. 理解 XML 概念 2. DOM 对 XML 文档进行增删改查	1. 理解 XML 概念 2. 对 XML 文档进行增删改查	由教师示范、指导、学生自主学习	4
17	XML (二): DOM4J 对 XML 文档进行增删改查	1. 理解 XML 概念 2. DOM4J 对 XML 文档进行增删改查	1. 理解 XML 概念 2. 对 XML 文档进行增删改查	由教师示范、指导学生自主学习	4
18	嗖嗖移动业务大厅项目	1. 理解嗖嗖业务大厅项目业务需求。 2. 编写项目代码。	3. 理解嗖嗖业务大厅项目业务需求。 编写项目代码。	由教师示范、指导学生自主学习	4

## 六、课程实施建议

(一) 教学建议 (从教学条件、教学方法与手段、课程资源的开发与利用、教材选用等方面进行说明)

### 1. 教学资料开发建议

- (1) 结合实际软件开发工程案例, 进行课件制作、教学环境模拟。
- (2) 与企业专家合作编写符合技能培养的实验实训教材, 是学生学有所用。
- (3) 提供网络资源学习平台, 使学生不拘泥于固定的时间、地点学习, 提高学生学习的灵活性。

### 2. 教学资源使用建议

(1) 学习包和教材, 都是素材性的课程资源, 但教材是知识的载体, 而学习包是引导学生学习的载体。因此, 要开发学习包, 突出以学生为中心的学习过程, 将本课程的教学内容与地区经济的发展密切结合, 并体现学生职业生涯发展的需要。

(2) 教学中学生的问题、困惑、见解、情感和体验等都是动态生成的课程资源, 老师应重视这些教学过程中动态生成的课程资源, 并主动性和创造性地运用, 使以项目为主的教学充满生机与活力。

(3) 本课程的实施, 不仅应充分开发校内的课程资源, 还应利用校外的课程资源, 如相关企业的岗位实习和实践, 尽量做到校企结合、工学结合。

(4) 积极开发教学资源: 课程标准、实训指导书、授课计划等教学文件, 以及课件、习题、案例库、网络方案、工具软件和网络资源等。

充分利用网络资源，了解最新的技术策略和实施方式，培养学生自主学习的能力。

### 3. 教材

选用教材：《Java 语言程序设计》 王新萍 电子工业出版社

#### (二) 考核建议

考核主要由平时成绩、理论考试以及实训项目成绩组成。

##### 1. 平时成绩

平时成绩强调平时的出勤、课堂纪律、课堂表现和作业，教师通过千分制的实施可以在整个授课过程中监控学生的学习效果以及学习态度等等。平时成绩占全部考试的 30%。

##### 2. 理论考试

理论考试采用传统的笔试方式，为了更好的达到考核的目的，我们采用第三方出题的形式来组织理论考试。理论考试占全部考试的 40%。

##### 3. 项目评审

实施项目内部考核及教师考核相结合的原则，在项目实训末期，要进行项目评审，根据项目的完成情况给学生打分。项目评审成绩占全部成绩的 30%（包括学生自我评价及对同伴的评价、完成项目效果评价以及教师评价，其中自我评价(占 30%)、项目效果评价(占 30%)、教师评价(占 40%)。）

## 七、需要说明的其他问题（参考资料、所需仪器、设备、教学软件等）

### 1. 参考资料

(1) 《Java 程序设计项目教程》 成奋华高等教育出版社

(2) 《Java 程序设计案例教程》钱银中编著机械工业出版社(国家级精品课程配套教材 21 世纪高职高专系列教材)

(3) 《Java 程序设计实例教程》马世霞机械工业出版社

(4) 与 Java 程序设计相关的其它书籍、网络资源

2. 硬件要求：微机、网络。

3. 软件要求：JDK、SQL Server、idea、MyEclipse 等软件。

## 《MySql 数据库》课程标准

## 一、课程基本信息

课程名称	MySql 数据库				
课程代码		学时	72	学分	4
授课时间	第 3 学期		适用专业	大数据技术与应用	
课程性质	岗位能力课程				
先修课程	Java 程序设计		后续课程	Java Web 应用开发	

## 二、课程定位

本课程是计算机应用技术专业的岗位能力课程，目的是通过课堂教学和实践教学相结合，使学生能够掌握 MySQL 的安装与配置、数据库的创建、数据表的创建、数据类型和运算符、MySQL 函数、查询数据、数据表的操作（插入、更新与删除数据）、索引、存储过程和函数、视图、触发器、用户管理、数据备份与还原、日志以及性能优化。通过 1 个综合案例的数据库设计，让学生进一步了解 MySQL 在实际工作中的应用。

## 三、课程设计思路

本课程主要以教材为参考，融入经典的项目案例，同时借鉴了一些企业级项目模块的雏形，有利于学生的理解与学习。不失传统性，任何一门语言课程都要了解该语言的书写风格与运行的平台特点，但是在学习过程中，又主要通过项目实例让学生进一步学习基础知识同时也见到了项目雏形。通过前期近七十多课时的学习，总体上掌握该门课程的特点与学习思路，然后利用较综合性项目培养学生对软件组织结构、思想、过程的设计。新课程理念要求充分的调动学生思考和动手能力，先学后教，精讲多练。

## 四、课程目标

## (一) 能力目标

1. 培养学生规范编码和良好的程序设计风格；
2. 培养学生发现问题、分析问题和解决问题的能力。

## (二) 知识目标

1. 掌握数据库的基本概念
2. 掌握安装并配置 MySql 数据库的方法
3. 掌握使用 SQLyog 管理工具连接并操作 MySql 数据库
4. 掌握使用 SQL 语句操作数据库
5. 能够使用数据库实现数据的增删改查
6. 掌握分页查询

7. 能够使用事务插入多条记录
8. 能够使用视图查看记录
9. 掌握数据表索引的创建
10. 掌握数据库的备份与恢复
11. 了解数据库规范

### (三) 素质目标

1. 培养按时完成项目的良好习惯，增强集体和社会责任感；
2. 培养团队协作和良好的沟通能力；
3. 培养学生务实、创新、诚实、守信的性格。

## 五、课程内容及要求

序号	教学内容	能力目标	知识目标	教学方法及手段	学时
1	数据库相关术语; 常见数据库; 数据库基本概念	能够初步认识数据库	掌握数据库的基本概念	由教师阐述, 说明	4
2	安装配置 MySQL 数据库; 演示 SQLyog 的使用方法, SQL 语句	能够自主安装配置 MySQL 数据库; 自主安装并简单使用 SQLyog 工具; 初步掌握 SQL 语句	1. 掌握 MySQL 以及 SQLyog 的安装配置 2. 掌握本章的 SQL 语句	由教师示范、指导, 学生小组协作	4
3	MySQL 管理	能够在 windows 系统和 linux 系统中分别启动和关闭 MySQL 服务, 可以添加 MySQL 用户,	1. 启动及关闭 MySQL 服务器 2. MySQL 用户设置	由教师示范、指导, 学生小组协作	4
4	创建数据库	能够分别使用命令和 SQLyog 工具创建数据库、删除数据库	1. 数据库的创建 2. 数据库的删除	由教师示范、指导, 学生小组协作	4
5	MySQL 数据类型	能够了解 MySQL 常用的数据类型, 并可以根据数据内容选择最优数据类型	1. 数值类型 2. 日期类型 3. 字符串类型	由教师示范、指导, 学生小组协作	4
6	MySQL 创建数据表	能够使用 SQL 语句创建数据表, 并插入正确的数据	1. 通过命令提示符创建 2. 通过 SQLyog 工具手动创建	由教师示范、指导, 学生小组协作	4
7	MySQL 常用存储引擎; 增删改查常用语句; 简单子查询; IN 子查询	实现数据库的增删改查	1. 掌握数据表增删改查语句的编写 2. 掌握简单的子查询 3. 掌握 IN 子查询	由教师示范、指导、学生小组协作完成	12
8	EXISTS 子查询; GROUP BY 子句; LIMIT 子句; SQL 语句综合查询	能够完成复杂的查询案例	1. 掌握分组查询 2. 掌握左右外三种连接	由教师示范、指导, 学生小组协作	8



9	MySQL 的事务、视图、索引、备份和恢复	使用事务插入多条数据，使用视图查看特定数据，创建数据表索引，数据库的备份和恢复	1. 会使用事务保证操作数据的完整性 2. 会创建并使用视图 3. 会创建并使用索引 4. 掌握数据库的备份和恢复	由教师示范、指导，学生小组协作	8
10	数据库规范化	完成指定数据库设计	1. 了解数据库设计的步骤 2. 掌握如何绘制数据库的 E-R 图 3. 掌握如何绘制数据库的模型图 4. 会使用三大范式实现数据库设计规范化	由教师示范、指导，学生小组协作	8
11	综合实战-银行 ATM 存取款机系统	能够按要求完成遵循三大范式的数据库的设计，并添加约束，对合法的业务数据进行增删改查操作，能够使用事务、视图实现业务处理	1. 掌握 MySQL 的用户管理 2. 会使用 SQL 语句操作数据 3. 会使用事务保证数据完整性 4. 会创建并使用视图 5. 会创建并使用索引 6. 掌握数据库的备份和恢复	由教师示范、指导，学生小组协作	12

## 六、课程实施建议

(一) 教学建议（从教学条件、教学方法与手段、课程资源的开发与利用、教材选用等方面进行说明）

### 1. 教学资料开发建议

- (1) 结合实际软件开发工程案例，进行课件制作、教学环境模拟。
- (2) 与企业专家合作编写符合技能培养的实验实训教材，是学生学有所用。
- (3) 提供网络资源学习平台，使学生不拘泥于固定的时间、地点学习，提高学生学习的灵活性。

### 2. 教学资源使用建议

(1) 学习包和教材，都是素材性的课程资源，但教材是知识的载体，而学习包是引导学生学习的载体。因此，要开发学习包，突出以学生为中心的学习过程，将本课程的教学内容与地区经济的发展密切结合，并体现学生职业生涯发展的需要。

(2) 教学中学生的问题、困惑、见解、情感和体验等都是动态生成的课程资源，老师应重视这些教学过程中动态生成的课程资源，并主动性和创造性地运用，使以项目为主的教学充满生机与活力。

(3) 本课程的实施，不仅应充分开发校内的课程资源，还应利用校外的课程资源，如相关企业的岗位实习和实践，尽量做到校企结合、工学结合。

(4) 积极开发教学资源：课程标准、实训指导书、授课计划等教学文件，以及课件、习题、案例库、网络方案、工具软件和网络资源等。

充分利用网络资源，了解最新的技术策略和实施方式，培养学生自主学习的能力。

### 3. 教材

选用教材：《高性能 MySQL》（第 3 版）[美] 保罗·迪布瓦（Paul DuBois）人民邮电出版社

## （二）考核建议

考核主要由平时成绩、理论考试以及实训项目成绩组成。

### 1. 平时成绩

平时成绩强调平时的出勤、课堂纪律、课堂表现和作业，教师通过千分制的实施可以在整个授课过程中监控学生的学习效果以及学习态度等等。平时成绩占全部考试的 30%。

### 2. 理论考试

理论考试采用传统的笔试方式，为了更好的达到考核的目的，我们采用第三方出题的形式来组织理论考试。理论考试占全部考试的 40%。

### 3. 项目评审

实施项目内部考核及教师考核相结合的原则，在项目实训末期，要进行项目评审，根据项目的完成情况给学生打分。项目评审成绩占全部成绩的 30%（包括学生自我评价及对同伴的评价、完成项目效果评价以及教师评价，其中自我评价（占 30%）、项目效果评价（占 30%）、教师评价（占 40%）。）

## 七、需要说明的其他问题（参考资料、所需仪器、设备、教学软件等）

### 1. 参考资料

（1）《MySQL 必知必会》（英）福塔 著，刘晓霞，钟鸣 译 人民邮电出版社

（2）《高性能 MySQL》（第 3 版）[美] 保罗·迪布瓦（Paul DuBois）人民邮电出版社

（3）《MySQL 管理之道：性能调优、高可用与监控》（第 2 版）贺春旻 机械工业出版社

（4）与 MySQL 数据库相关的其它书籍、网络资源

2. 硬件要求：微机、网络。

3. 软件要求：MySQL、SQLyog 等软件。

## 《Python 程序设计》课程标准

## 一、课程基本信息

课程名称	Python 程序设计				
课程代码		学时	72	学分	4
授课时间	第三学期	适用专业	大数据技术与应用专业		
课程性质	岗位能力课程				
先修课程	Java 程序设计	后续课程	Hadoop 系统基础		

## 二、课程定位

本课程是大数据技术与应用专业的岗位能力课程，目的是通过课堂教学和实践教学相结合，使学生能够深入理解 Python 及面向对象概念，清楚的了解 Python 软件开发工作流程，建立起应用程序的概念，最终能够掌握 Python 软件开发的基本方法、基本技能，培养利用 Pygame 等库进行 Python 应用软件产品的分析、设计、编码、测试的综合应用能力，培养科学的思维方法，灵活运用知识的能力，实验操作能力，使学生具有较强的发现问题、分析问题、解决问题的能力，具有毕业后的直接上岗能力（或经短期培训后上岗），并且毕业时已具备一定的 Python 软件开发经验。

## 三、课程设计思路

本课程主要以教材为参考，融入经典的项目案例，同时借鉴了一些企业级项目模块的雏形，有利于学生的理解与学习。不失传统性，任何一门语言课程都要了解该语言的书写风格与运行的平台特点，但是在学习过程中，又主要通过项目实例让学生进一步学习基础知识同时也见到了项目雏形。通过前期近一百多课时的学习，总体上掌握该门课程的特点与学习思路，然后利用较综合性项目培养学生对软件组织结构、思想、过程的设计。新课程理念要求充分的调动学生思考和动手能力，先学后教，精讲多练。

## 四、课程目标

## （一）能力目标

## （一）能力目标

1. 培养学生规范编码和良好的程序设计风格；
2. 培养学生面向对象编程的思维和提高逻辑思维能力；

3. 培养学生发现问题、分析问题和解决问题的能力。

(二) 知识目标

1. 掌握 Python 语言基础;
2. 熟练掌握对象和类;
3. 掌握继承与多态;
4. 熟练掌握数组和字符串;
5. 掌握 Python 的异常处理;
6. 掌握 Python 的输入/输出;
7. 了解 GUI 程序设计;
8. 熟练掌握 Python 的集合;
9. 了解面向对象的常用设计模式;
10. 掌握 Python 程序设计的思想和方法。

(三) 素质目标

1. 培养按时完成项目的良好习惯，增强集体和社会责任感;
2. 培养团队协作和良好的沟通能力;
3. 培养学生务实、创新、诚实、守信的性格。

五、课程内容及要求

序号	教学内容	能力目标	知识目标	教学方法及手段	学时
1	计算机和程序	了解计算机软硬件及 Python 程序内容	<ol style="list-style-type: none"> <li>1. 了解计算系统中硬件和软件各自的作用。</li> <li>2. 学习计算机科学家研究的领域和他们使用的技术。</li> <li>3. 了解现代计算机的基本设计。</li> <li>3. 了解计算机编程语言的形式和功能。</li> <li>4. 开始使用 Python 编程语言。</li> </ol>	由教师示范、指导、讲解	4
2	编写简单程序	能够熟练应用输入输出赋值及循环语句完成简单程序编写。	<ol style="list-style-type: none"> <li>1. 知道有序的软件开发过程的步骤。</li> <li>2. 了解遵循输入、处理、输出(IPO)模式的程序,并能够以简单的方式修改它们。</li> <li>3. 了解构成有效 Python 标识符和表达式的规则。</li> <li>4. 能够理解和编写 Python 语句,将信息输出到屏幕,为变量赋值,获取通过键盘输入的信息,并执行计数循环。</li> </ol>	由教师示范、指导,学生小组协作	4
3	数字计算	掌握数字计算	<ol style="list-style-type: none"> <li>1. 理解数据类型的概念。</li> <li>2. 熟悉 Python 中的基本数值数据类型。</li> <li>3. 理解数字在计算机上如何表示</li> </ol>	由教师示范、指导、学生小组协作完成	4

			<p>的基本原理。</p> <p>4. 能够使用 Python 的 math 库。</p> <p>5. 理解累积器程序模式。</p> <p>6. 能够阅读和编写处理数值数据的程序。</p>		
4	对象和图	掌握图形及 graphics 模块	<p>1. 理解对象的概念以及如何用它们来简化编程。</p> <p>2. 熟悉 graphics 库中可用的各种对象。</p> <p>3. 能够在程序中创建对象并调用适当的方法来进行图形计算。</p> <p>4. 了解计算机图形学的基本概念,特别是坐标系统和坐标变换的作用。</p> <p>5. 了解如何在图形编程语境中使用基于鼠标和基于文本的输入。</p> <p>6. 能够使用 graphics 库编写简单的交互式图形程序。</p>	由教师示范、指导、学生独立编程实现点、圆、圆柱类的设计与实现	4
5	序列、字符串、列表和文	掌握序列、字符串、列表和文件的操作	<p>1. 了解字符串数据类型以及如何在计算机中表示字符串。</p> <p>2. 熟悉通过内置函数和字符串方法对字符串执行的各种操作。</p> <p>3. 理解序列和索引的基本概念,因为它们适用于 Python 的字符串和列表。</p> <p>4. 能够用字符串格式化来产生有吸引力的、富含信息的程序输出。</p> <p>5. 了解在 Python 中读取和写入文本文件的基本文件处理概念和技术。</p> <p>6. 了解加密的基本概念。</p> <p>7. 理解和编写处理文本信息的程序。</p>	由教师示范、指导、学生独立编程实现图形类的设计与实现	4
6	定义函数	掌握函数的编写	<p>1. 了解程序员为什么将程序分成多组合作的函数。</p> <p>2. 能够在 Python 中定义新的函数。</p> <p>3. 理解 Python 中函数调用和参数传递的细节。</p> <p>4. 利用函数来编程,减少代码重复并增加程序的模块性。</p>	由教师示范、指导、学生小组协作完成图形类的设计与实现	4
7	判断结构	掌握判断语句	<p>1. 利用 Python 的 if 语句来理解简单的判断编程模式及其实现。</p> <p>2. 利用 Python 的 if-else 语句来理解两路判断编程模式及其实现。</p> <p>3. 利用 Python 的 if-elif-else 语句来理解多路判断编程模式及其实现。</p> <p>4. 理解异常处理的思想,并能够编写简单异常处理代码,捕捉标准的 Python 运行时错误。</p> <p>5. 理解布尔表达式和布尔数据类型的概念。</p> <p>6. 能够阅读、编写和实现使用判断结构的算法,包括使用系列判断和嵌套判断结构的算法。</p>	由教师示范、指导、学生自主学习	4

8	循环结构和布尔	掌握循环结构和布尔值	<ol style="list-style-type: none"> <li>1. 理解确定和不定循环的概念,以及它们用 Python 的 for 和 while 语句的实现。</li> <li>2. 理解交互式循环和哨兵循环的编程模式,以及它们用 Python 的 while 语句的实现。</li> <li>3. 理解文件结束循环的编程模式,以及在 Python 中实现这种循环的方法。</li> <li>4. 能为涉及循环模式(包括嵌套循环结构)的问题设计和实现解决方案。</li> <li>5. 理解布尔代数的基本思想,并能分析和编写涉及布尔运算符的布尔表达式。</li> </ol>	由教师示范、指导、学生自主学习	4
9	模拟与设计	理解并掌握模拟与设计	<ol style="list-style-type: none"> <li>1. 理解模拟的应用可能是解决现实问题的一种方式。</li> <li>2. 理解伪随机数及其在蒙特卡罗模拟中的应用。</li> <li>3. 理解并能应用自顶向下和螺旋式设计技术来编写复杂的程序。</li> <li>4. 理解单元测试,并能将这种技术应用于复杂程序的实现和调试。</li> </ol>	由教师示范、指导,学生自主学习认识常用组件	4
10	定义类	掌握类、对象、封装和控件	<ol style="list-style-type: none"> <li>1. 领会定义新类如何能为复杂程序提供结构。</li> <li>2. 能够阅读并编写 Python 类定义。</li> <li>3. 理解封装的概念,以及它如何有助于构建模块化的、可维护的程序。</li> <li>4. 能够编写包含简单类定义的程序。</li> <li>5. 能够编写包含创新(程序员设计的)控件的交互式图形程序。</li> </ol>	由教师示范、指导,学生小组学习认识容器与菜单	4
11	数据集合	掌握数组的使用	<ol style="list-style-type: none"> <li>1. 了解使用列表(数组)来表示相关数据的集合。</li> <li>2. 熟悉用于操作 Python 列表的函数和方法。</li> <li>3. 能够编程用列表管理信息集合。</li> <li>4. 能够编程利用列表和类来构造复杂数据。</li> <li>5. 了解用 Python 字典存储无顺序集合。</li> </ol>	由教师示范、指导学生小组学习	4
12	面向对象设计	掌握面向对象思想及操作	<ol style="list-style-type: none"> <li>1. 理解面向对象设计的过程。</li> <li>2. 能够阅读和理解面向对象的程序。</li> <li>3. 理解封装、多态和继承的概念,因为它们从属于面向对象的设计和编程。</li> <li>4. 能够利用面向对象设计来设计中等复杂程度的软件。</li> </ol>	由教师示范、指导学生小组学习	4
13	算法设计与递归	掌握查找、排序及递归算法	<ol style="list-style-type: none"> <li>1. 理解分析算法效率的基本技巧。</li> <li>2. 知道查找是什么,并且理解线性和二分查找的算法。</li> <li>3. 理解递归定义和函数的基本原理,并能够编写简单的递归函数。</li> <li>4. 深入理解排序,并理解选择排序</li> </ol>	由教师示范、指导学生小组学习完成画布上画线实训	4

			和归并排序的算法。		
14	Gui 的实现—Tkinter	掌握 Tkinter	1. 利用 kinter 实现 GUI	由教师示范、指导学生自主学习	8
15	综合实例：Pygame 游戏开发	掌握 Pygame	2. 利用游戏复习巩固前面的内容	由教师示范、指导学生自主学习	12

## 六、课程实施建议

(一) 教学建议（从教学条件、教学方法与手段、课程资源的开发与利用、教材选用等方面进行说明）

### 1. 教学资料开发建议

- (1) 结合实际软件开发工程案例，进行课件制作、教学环境模拟。
- (2) 与企业专家合作编写符合技能培养的实验实训教材，是学生学有所用。
- (3) 提供网络资源学习平台，使学生不拘泥于固定的时间、地点学习，提高学生学习的灵活性。

### 2. 教学资源使用建议

(1) 学习包和教材，都是素材性的课程资源，但教材是知识的载体，而学习包是引导学生学习的载体。因此，要开发学习包，突出以学生为中心的学习过程，将本课程的教学内容与地区经济的发展密切结合，并体现学生职业生涯发展的需要。

(2) 教学中学生的问题、困惑、见解、情感和体验等都是动态生成的课程资源，老师应重视这些教学过程中动态生成的课程资源，并主动性和创造性地运用，使以项目为主的教学充满生机与活力。

(3) 本课程的实施，不仅应充分开发校内的课程资源，还应利用校外的课程资源，如相关企业的岗位实习和实践，尽量做到校企结合、工学结合。

(4) 积极开发教学资源：课程标准、实训指导书、授课计划等教学文件，以及课件、习题、案例库、网络方案、工具软件和网络资源等。

充分利用网络资源，了解最新的技术策略和实施方法，培养学生自主学习的能力。

### 3. 教材

选用教材：

《Python 语言程序设计》（第3版） 约翰·策勒 人民邮电出版社发行

### (二) 考核建议

考核主要由平时成绩、理论考试以及实训项目成绩组成。

#### 1. 平时成绩

平时成绩强调平时的出勤、课堂纪律、课堂表现和作业，教师通过千分制的实施可以在整个授课过程中监控学生的学习效果以及学习态度等等。平时成绩占全部考试的 30%。

#### 2. 理论考试

理论考试采用传统的笔试方式，为了更好的达到考核的目的，我们采用第三方出题的形式来组织理论考试。理论考试占全部考试的 40%。

#### 3. 项目评审

实施项目内部考核及教师考核相结合的原则，在项目实训末期，要进行项目评审，根据项目的完成情况给学生打分。项目评审成绩占全部成绩的 30%（包括学生自我评价及对同伴的评价、完成项目效果评价以及教师评价，其中自我评价（占 30%）、项目效果评价（占 30%）、教师评价（占 40%）。）

### 七、需要说明的其他问题（参考资料、所需仪器、设备、教学软件等）

#### 1. 参考资料

- (1) 《Python 程序设计》（第 2 版） 清华大学出版社
- (2) 《从零基础入门学习 Python》 清华大学出版社
- (3) 《Python 程序设计基础》 人民邮电出版社
- (4) 与 Python 程序设计相关的其它书籍、网络资源

#### 2. 硬件要求：微机、网络。

#### 3. 软件要求：Vmware workstaion 15 pro、Linux、Python3.5 等软件。



## 《Hadoop 系统基础》课程标准

## 一、课程基本信息

课程名称	Hadoop 系统基础				
课程代码		学时	108	学分	6
授课时间	第三学期	适用专业	大数据技术与应用		
课程性质	岗位能力课程				
先修课程	Java 程序设计, MySQL 数据库	后续课程	HBase 系统基础, Hive 数据仓库系统		

## 二、课程定位

本课程是计算机应用技术专业的岗位能力课程,目的是通过课堂教学和实践教学相结合,使学生能够深入理解

以 Hadoop 和 Spark 为核心,阐述了基本这两种通用大数据处理平台的应用开发技术。在 Hadoop 生态圈中,从 HDFS 初识分布式存储系统;以 MapReduce 详解分布式计算的步骤;让我们深入了解大数据处理平台的应用开发,结合实践,具有毕业后的直接上岗能力,并且毕业时已具备一定的大数据处理平台的开发经验。

## 三、课程设计思路

本课程主要以教材为参考,融入经典的项目案例,同时借鉴了一些企业级项目模块的雏形,有利于学生的理解与学习。不失传统性,任何一门语言课程都要了解该语言的书写风格与运行的平台特点,但是在学习过程中,又主要通过项目实例让学生进一步学习基础知识同时也见到了项目雏形。通过前期近一百多课时的学习,总体上掌握该门课程的特点与学习思路,然后利用较综合性项目培养学生对软件组织结构、思想、过程的设计。新课程理念要求充分的调动学生思考和动手能力,先学后教,精讲多练。

## 四、课程目标

## (一) 能力目标

## (一) 能力目标

1. 培养学生规范编码和良好的程序设计风格;
2. 培养学生面向对象编程的思维和提高逻辑思维能力;
3. 培养学生发现问题、分析问题和解决问题的能力。

(二) 知识目标

1. 了解大数据基本概念；
2. 了解 Hadoop；
3. 了解 HDFS；
4. 了解 HDFS 运行原理。
5. 了解 HDFS 高级知识；
6. 了解 MapReduce 编程模型；
7. 了解 YARN；
8. 了解 HDFS 新特性；
9. 了解 YARN 新特性；
10. 掌握安装 Hadoop 平台；
11. 掌握 HDFS 基本操作；
12. 掌握 MapReduce 应用开发；
13. 掌握 MapReduce 高级应用；

(三) 素质目标

1. 培养按时完成项目的良好习惯，增强集体和社会责任感；
2. 培养团队协作和良好的沟通能力；
3. 培养学生务实、创新、诚实、守信的性格。

五、课程内容及要求

序号	教学内容	能力目标	知识目标	教学方法及手段	学时
1	初始大数据：大数据基本概念，大数据带来的挑战	掌握大数据的基本概念和基本特性，以及大数据带给社会，企业的挑战有哪些	1. 了解大数据的基本概念和基本特性。 2. 了解大数据带来的挑战，大数据对现有数据库的挑战，实时性的技术挑战，对数据中心，运维团队的挑战	由教师示范、指导学生小组学习	8
2	初始 Hadoop：Hadoop 概述，Hadoop 生态圈，Hadoop 应用案例	掌握 Hadoop 的基本概念，掌握 Hadoop 的核心构成，了解 Hadoop 生态圈中各个组件的功能	1. 了解 Hadoop 的概念以及 Hadoop 发展史 2. 掌握 Hadoop 框架构成，分布式文件系统，分布式计算系统，资源管理系统 3. 掌握 Hadoop 生	由教师示范、指导学生小组学习	8

			态圈		
3	安装 Hadoop 平台：安装虚拟机，安装 Linux 系统，安装 Hadoop 伪分布式环境	掌握虚拟机，centOS, hadoop 的安装	1. 了解虚拟机概述 2. 掌握虚拟机的安装方式，掌握 Linux 系统的安装方式 3. 安装 Hadoop 伪分布式环境	由教师示范、指导学生小组学习	8
4	HDFS 入门：认识 HDFS，HDFS 基础，HDFS 架构	掌握 HDFS 的基本概念与优缺点 掌握 HDFS 架构中的各个组件功能	1. 了解 HDFS 产生背景以及简介 2. 了解 HDFS 的优缺点 3. 掌握 HDFS 架构中的各个组件功能	由教师示范、指导、学生独立编程实现点 HDFS 的安装	8
5	HDFS 基本操作：使用 HDFS shell 访问，使用 Java API 访问	掌握使用 HDFS shell 访问命令 掌握使用 java API 来操作 HDFS 上的数据	1. 掌握 HDFS shell 的方式来访问 HDFS 上的数据 2. 掌握使用 java API 操作 HDFS 的常用操作	由教师示范、指导、学生独立编程实现图形类的设计与实现	8
6	HDFS 运行原理：HDFS 读写流程，HDFS 副本机制，HDFS 负载均衡，HDFS 机架感知	掌握 HDFS 文件读写流程 了解 HDFS 的副本机制 了解 HDFS 文件数据的负载均衡和机架感知	1. 掌握 HDFS 文件的读写流程 2. 了解 HDFS 的副本机制以及容错机制。 3. 了解 HDFS 负载均衡以及机架感知	由教师示范、指导、学生小组协作完成分布式文件系统 HDFS 的读写流程	8
7	HDFS 高级知识：Hadoop 序列化机制，SequenceFile, MapFile	掌握 Hadoop 的序列化 掌握 SequenceFile 特点以及读操作 掌握 MapFile 写操作	1. 了解序列化与反序列化的概念 2. 掌握 Hadoop 的序列化 3. 掌握 SequenceFile 概述与特点以及读操作 MapFile 写操作	由教师示范、指导、学生自主学习，并完成序列化对象，并将序列化的对象反序列化出来	8
8	认识 MapReduce 编程模型：MapReduce 基础，MapReduce 编程模型，MapReduce 词频统计编程实例	掌握 MapReduce 的设计目标，了解 MapReduce 的特点与 MapReduce 编程模型。	1. 了解 MapReduce 概念。 2. 熟练掌握 MapReduce 的设计目标 3. 了解 MapReduce 编程模型	由教师示范、指导、学生自主学习 MapReduce 的编制与运行	8
9	MapReduce 应用开发：MapReduce 输入/输出	了解 MapReduce 类型概述。	1. 了解 MapReduce 类型	由教师示范、指导、学生自主学习	10

	类型, MapReduce 输入格式, MapReduce 输出格式, Combiner 操作, Partitioner 操作, 自定义 RecordReader	熟练掌握 MapReduce 输出、输入格式。实现 Combiner 操作, 与 Partitioner 操作。	输入和输出 key/value 的概念。2. 掌握 MapReduce 输入格式、会使用 InputFormat 接口。3. 实现 Combiner 操作。	MapReduce 的输入与输出, 多加练习 Combiner 的操作原理, 理解定义 RecordReader。	
10	MapReduce 高级应用: 使用 MapReduce 实现 join 操作, 使用 MapReduce 实现排序, 使用 MapReduce 实现二次排序, 使用 MapReduce 合并小文件	掌握 MapReduce 实现 join 操作。掌握了 MapReduce 排序的实现原理。	1. 熟练掌握 MapReduce 实现 join 操作。 2. 实现 MapReduce 排序, 理解原理。	由教师示范、指导、学生自主学习 MapReduce 实现 join 操作, 多加练习 MapReduce 排序, 理解其原理。	8
11	初始 YARN: YARN 产生背景, YARN 简介, YARN 架构设计	了解 YARN; 背景, 简介, 架构设计	了解 YARN; 背景, 简介, 架构设计	由教师示范、指导学生小组学习了解 YARN 是什么东西, 理解其含义, 以及应用场景	8
12	了解 HDFS 新特性: HDFS NameNode 高可用机制, HDFS NameNode Federation, HDFS REST API, DistCp 工具	了解 HDFS 新特性、HDFS NameNode 高可用机制。熟练掌握 HDFS NameNode Federation, HDFS REST API, DistCp 工具	了解 HDFS 新特性、HDFS NameNode 高可用机制。熟练掌握 HDFS NameNode Federation, HDFS REST API, DistCp 工具	由教师示范、指导学生小组学习明白 HDFS 新特性、HDFS NameNode 高可用机制。熟练掌握课上所讲的一些工具。	10
13	了解 YARN 新特性: ResourceManager 自动重启, ResourceManager 高可用机制	了解 YARN 新特性, ResourceManager 自动重启概念、ResourceManager 高可用机制	了解 YARN 新特性, ResourceManager 自动重启概念、ResourceManager 高可用机制	由教师示范、指导学生小组学习 YARN 新特性。ResourceManager 自动重启概念、明白高可用机制。	8

## 六、课程实施建议

(一) 教学建议 (从教学条件、教学方法与手段、课程资源的开发与利用、教材选用等方面进行说明)

### 1. 教学资料开发建议

- (1) 结合实际软件开发工程案例, 进行课件制作、教学环境模拟。
- (2) 与企业专家合作编写符合技能培养的实验实训教材, 是学生学有所用。
- (3) 提供网络资源学习平台, 使学生不拘泥于固定的时间、地点学习, 提高学生学习的灵活性。

### 2. 教学资源使用建议

- (1) 学习包和教材, 都是素材性的课程资源, 但教材是知识的载体, 而学习包是引导学生学习的载体。因此, 要开发学习包, 突出以学生为中心的学习过

程，将本课程的教学内容与地区经济的发展密切结合，并体现学生职业生涯发展的需要。

(2) 教学中学生的问题、困惑、见解、情感和体验等都是动态生成的课程资源，老师应重视这些教学过程中动态生成的课程资源，并主动性和创造性地运用，使以项目为主的教学充满生机与活力。

(3) 本课程的实施，不仅应充分开发校内的课程资源，还应利用校外的课程资源，如相关企业的岗位实习和实践，尽量做到校企结合、工学结合。

(4) 积极开发教学资源：课程标准、实训指导书、授课计划等教学文件，以及课件、习题、案例库、网络方案、工具软件和网络资源等。

充分利用网络资源，了解最新的技术策略和实施方法，培养学生自主学习的能力。

### 3. 教材

选用教材：《Hadoop 系统基础》

#### (二) 考核建议

考核主要由平时成绩、理论考试以及实训项目成绩组成。

##### 1. 平时成绩

平时成绩强调平时的出勤、课堂纪律、课堂表现和作业，教师通过千分制的实施可以在整个授课过程中监控学生的学习效果以及学习态度等等。平时成绩占全部考试的 30%。

##### 2. 理论考试

理论考试采用传统的笔试方式，为了更好的达到考核的目的，我们采用第三方出题的形式来组织理论考试。理论考试占全部考试的 40%。

##### 3. 项目评审

实施项目内部考核及教师考核相结合的原则，在项目实训末期，要进行项目评审，根据项目的完成情况给学生打分。项目评审成绩占全部成绩的 30%（包括学生自我评价及对同伴的评价、完成项目效果评价以及教师评价，其中自我评价（占 30%）、项目效果评价（占 30%）、教师评价（占 40%）。）

## 七、需要说明的其他问题（参考资料、所需仪器、设备、教学软件等）

### 1. 参考资料

(1) 《Hadoop 实战 2》作者：陆嘉恒

- (2) 《Hadoop 权威指南》
  - (3) 《高可用的 HDFS》 - Hadoop 分布式文件系统深度实践
2. 硬件要求：微机、网络。
  3. 软件要求：JDK、Hadoop、Linux、虚拟机等软件。

## 《HBase 系统基础》课程标准

## 一、课程基本信息

课程名称	HBase 系统基础				
课程代码		学时	72	学分	4
授课时间	第四学期	适用专业	大数据技术与应用		
课程性质	岗位能力课程				
先修课程	Java 程序设计, Hadoop 系统基础	后续课程	Spark 大数据技术与应用		

## 二、课程定位

HBase 是一个分布式的、面向列的开源数据库，不同于一般的关系数据库，它是一个适合于非结构化数据存储的数据库。本课程主要讲述了 HBase 详细的架构原理及特点、HBase 内部各个角色的详细介绍、安装配置、HBase 的 Shell 操作、新旧版本的读写数据详细流程、HBase 的 API 操作。本课程是计算机应用技术专业的岗位能力课程，目的是通过课堂教学和实践教学相结合，使学生能够深入理解 HBase 的概念、原理、服务架构，最终能够掌握使用 MapReduce 以及 Hive 对 HBase 数据分析、Rowkey 设计、预分区设计、调优策略安装部署、集群配置及 web 环境搭建及综合应用的方法。

## 三、课程设计思路

最近几年的大数据是特别火爆，国家也在大力发展大数据项目发展。未来，大数据技术在 IT 发展中，具有重要的战略意义。本课程主要以教材为参考，融入经典的项目案例，同时借鉴了一些企业级项目模块的雏形，有利于学生的理解与学习。不失传统性，任何一门语言课程都要了解该语言的书写风格与运行的平台特点，但是在学习过程中，又主要通过项目实例让学生进一步学习基础知识同时也见到了项目雏形。本课程分为三个大的部分对 HBase 进行系统化的介绍，由浅入深，最后以一个实践的项目为例说明怎么使用 HBase 以及怎样的设计表能够提高查询性能，总体上掌握该门课程的特点与学习思路。

## 四、课程目标

## (一) 能力目标

1. 培养学生规范的设计表、建立表的预分区的能力；
2. 培养学生调优策略安装部署、集群配置及 web 环境搭建的能力；

(二) 知识目标

1. 掌握 HBase 应用场景、环境搭建；
2. 了解与 RDBMS 比较的优缺点；
3. 掌握表的模型；
4. 熟练 HBase 架构；
5. 了解数据存储模型；
6. 掌握 HBase Java API 使用；
7. 掌握 HBase 与 MapReduce 集成；
8. 熟练掌握 Hbase 部署环境准备 ；
9. 掌握 Hbase 伪分布式的启动及 hbase 命令的使用；
10. 掌握 Hbase shell 中表的 DDL 操作；
11. 掌握 Hbase 的读写流程；
12. 掌握 Hbase 表的设计 。

(三) 素质目标

1. 培养按时完成项目的良好习惯，增强集体和社会责任感；
2. 培养团队协作和良好的沟通能力；
3. 培养学生务实、创新、诚实、守信的性格。

五、课程内容及要求

序号	教学内容	能力目标	知识目标	教学方法及手段	学时
1	HBase 基本使用及存储设计	掌握 HBase 基本概念、掌握 HBase 环境搭建	1. HBase 的介绍及发展 2. HBase 与 MYSQL 的存储比较 3. Hbase 部署环境准备	由教师示范、指导学生小组学习	8
2	Hbase 伪分布式配置	了解所有进程运行在同一个节点上,不同进程运行在不同的 JVM 当中	1. Hbase 伪分布式配置文件的修改 2. Hbase 伪分布式的启动及 hbase 命令的使用	由教师示范、指导学生小组学习	8
3	Hbase shell 命令	掌握 Shell 基本命令,了解 HBase shell 中的帮助命令	1.namespace 的常用操作 2.表的 DDL 操作	由教师示范、指导学生小组学习	8
4	Hbase 的模型	掌握 HBase 物理模型、HBase 体系架构	1. Hbase 的物理模型 2. Hbase 的存储模型 3.HBase 的数据存储	由教师示范、指导学生小组学习	8
5	Hbase 的 namespace 表、meta 表及读写流程	掌握创建 namespace 删除 namespace 查看 namespace	1. Hbase 的读写流程 2. Hbase 中各组件的应用	由教师示范、指导学生小组学习	8



		列出全部 namespace			
6	Hbase Java API-环境配置	掌握配置 Hbase Java API 的基本步骤	1. Java API-通过 get 读取 2. Hbase 表中数据 3. Java API-put、delete 4. Java API-Scan 和过滤器	由教师示范、指导学生小组学习	8
7	Hbase 与 MapReduce 集成	掌握 HBase 与 MapReduce 的集成整合与常用操作	1. 环境变量的配置 2. rowcounter 测试 3 代码实现自定义需求	由教师示范、指导、学生自主学习 MapReduce 的编制与运行	10
8	案例：Hbase 的设计及企业优化	掌握基于具体的业务场景的实际动手操作	1. Hbase 表的设计 2. Hbase 性能优化-配置 snappy 压缩 3. PHoenix 的编译及安装部署	由教师示范、指导学生小组进行企业及项目的设计	14

## 六、课程实施建议

(一) 教学建议 (从教学条件、教学方法与手段、课程资源的开发与利用、教材选用等方面进行说明)

### 1. 教学资料开发建议

- (1) 结合实际软件开发工程案例, 进行课件制作、教学环境模拟。
- (2) 与企业专家合作编写符合技能培养的实验实训教材, 是学生学有所用。
- (3) 提供网络资源学习平台, 使学生不拘泥于固定的时间、地点学习, 提高学生学习的灵活性。

### 2. 教学资源使用建议

(1) 学习包和教材, 都是素材性的课程资源, 但教材是知识的载体, 而学习包是引导学生学习的载体。因此, 要开发学习包, 突出以学生为中心的学习过程, 将本课程的教学内容与地区经济的发展密切结合, 并体现学生职业生涯发展的需要。

(2) 教学中学生的问题、困惑、见解、情感和体验等都是动态生成的课程资源, 老师应重视这些教学过程中动态生成的课程资源, 并主动性和创造性地运用, 使以项目为主的教学充满生机与活力。

(3) 本课程的实施, 不仅应充分开发校内的课程资源, 还应利用校外的课程资源, 如相关企业的岗位实习和实践, 尽量做到校企结合、工学结合。

(4) 积极开发教学资源: 课程标准、实训指导书、授课计划等教学文件, 以及课件、习题、案例库、网络方案、工具软件和网络资源等。

充分利用网络资源, 了解最新的技术策略和实施方法, 培养学生自主学习的能力。

### 3. 教材

选用教材：《HBase 应用架构》 美 吉恩·马克斯帕加里 中国电力出版社

#### (二) 考核建议

考核主要由平时成绩、理论考试以及实训项目成绩组成。

##### 1. 平时成绩

平时成绩强调平时的出勤、课堂纪律、课堂表现和作业，教师通过千分制的实施可以在整个授课过程中监控学生的学习效果以及学习态度等等。平时成绩占全部考试的 30%。

##### 2. 理论考试

理论考试采用传统的笔试方式，为了更好的达到考核的目的，我们采用第三方出题的形式来组织理论考试。理论考试占全部考试的 40%。

##### 3. 项目评审

实施项目内部考核及教师考核相结合的原则，在项目实训末期，要进行项目评审，根据项目的完成情况给学生打分。项目评审成绩占全部成绩的 30%（包括学生自我评价及对同伴的评价、完成项目效果评价以及教师评价，其中自我评价（占 30%）、项目效果评价（占 30%）、教师评价（占 40%）。）

## 七、需要说明的其他问题（参考资料、所需仪器、设备、教学软件等）

### 1. 参考资料

- (1) 《HBase 权威指南》 人民邮电出版社（“十二五”国家重点图书出版规划项目）
- (2) 《HBase 实战》 人民邮电出版社
- (3) 《HBase 入门与实践》 人民邮电出版社
- (4) 《HBase 分布式存储系统应用》 水利水电出版社（普通高等教育新工科人才培养规划教材（大数据专业））

2. 硬件要求：微机、网络。

3. 软件要求：JDK、Hadoop、Linux、虚拟机等软件。

## 《Hive 数据仓库系统》课程标准

## 一、课程基本信息

课程名称	Hive 数据仓库系统				
课程代码		学时	72	学分	4
授课时间	第四学期	适用专业	大数据技术与技术专业		
课程性质	岗位能力课程				
先修课程	Hadoop 系统基础	后续课程	Spark 大数据技术与应用		

## 二、课程定位

本课程是计算机应用技术专业的岗位能力课程，目的是通过课堂教学和实践教学相结合，使学生能够深入理解 Hive，Hive 是一个构建在 Hadoop 上的数据仓库框架，是应 Facebook 每天生产的海量新兴社会网络数据进行管理和学习的需求而产生和发展的。解决 MapReduce 学习成本高、项目周期短、MapReduce 实现复杂查询逻辑开发难度大的问题。通过本课程学生学习，学生能够掌握如何完成大数据集的批处理作业。

## 三、课程设计思路

本课程旨在介绍如何使用 Hive 的 SQL 方法 HiveQL 来汇总、查询和分析存储在 Hadoop 分布式文件系统上的大数据集合。课程通过大量的实例，首先介绍如何在用户环境下安装和配置 Hive，并对 Hadoop 和 MapReduce 进行详尽阐述，终演示 Hive 如何在 Hadoop 生态系统进行工作。

## 四、课程目标

## (一) 能力目标

1. 培养学生规范编码和良好的程序设计风格；
2. 培养学生对大数据集分析处理的能力；
3. 培养学生发现问题、分析问题和解决问题的能力。

## (二) 知识目标

1. 掌握 Hive 的安装及配置
2. 熟练 Hive 命令
3. 掌握数据类型和文件格式
4. 熟练掌握 Hive 的数据定义

5. 掌握 Hive 的数据操作
6. 掌握 Hive 查询
7. 掌握 Hive 视图
8. 熟练 Hive 索引
9. 了解模式设计
10. 掌握 Hive 调优方法
11. 了解其他文件格式和压缩方法
12. 掌握 Hive 函数。

### (三) 素质目标

1. 培养按时完成项目的良好习惯，增强集体和社会责任感；
2. 培养团队协作和良好的沟通能力；
3. 培养学生务实、创新、诚实、守信的性格。

## 五、课程内容及要求

序号	教学内容	能力目标	知识目标	教学方法及手段	学时
1	基础知识： Hadoop 和 MapReduce 综述	熟悉 Hadoop 和 MapReduce 计算机模 型。 理解 MapReduce 的 基本原理。	1. 了解 haddop 和 MapReduce 综述 2. 掌握 MapReduce 的基 本原理	引导启发学生自动去学 习查找了解 Hadoop 的起 源及前期。	4
2	Hive 基本 操作	能够掌握如何安装 Hiew 的 Hadoop 集 群。	1. 掌握安装 Hive 的 Hadoop 集群。 2. 掌握 Hive 的启动方式 方 3. 掌握 Hive 命令	由教师示范、指导，学 生小组协作	6
3	数据类型 和文件格 式	掌握基本数据类型	1. 掌握 Hive 中的基本数 据类型 2. 集合数据类型 3. 掌握文本文件数据编 码 4. 了解读时模式	由教师示范、指导 、学生小组协作完成本 章的设计与实现	6
4	HiveQL:数 据定义	掌握 HiveQL 中数据 库和表的操作	1. 掌握数据库的创建和 修改。 2. 掌握表：管理表、外 部表、分区表等、删除表、 修改表	由教师示范、指导、学 生独立完成数据库及表 的操作。	8
5	HiveQL:数 据操作	掌握数据的操作	1. 掌握向管理表中装 载数据。 2. 通过查询语句向表 中插入数据。 3. 单个查询语句中创 建表并加载数据。 4. 导入导出数据	由教师示范、指导、学 生完成数据的操作	6
6	HiveQL:查 询	1. 掌握 HiveQL 查询	1. 掌握 SELECT...FROM... 语句 2. 掌握 WHERE 语句	由教师示范、指导 、学生小组协作完成 HiveQL 查询功能	6

			3. 掌握 GROUP BY 语句 4. 掌握 JOIN 语句等		
7	HiveQL: 视图	掌握 HiveQL 视图。	1. 能够使用视图来降低查询复杂度。2. 能够使用视图限制基于条件过滤数据。3. 能够通过动态分区中的视图和 map 类型。4	由教师示范、指导、学生自主学习	4
8	HiveQL: 索引	掌握 HiveQL 索引	1. 掌握 Hive 所有的创建、修改、显示及删除。	由教师示范、指导、学生自主学习 Hive 索引。	6
9	模式设计	了解 Hive 模式设计	1. 了解 Hive 设计的设计模式	由教师示范、指导，学生自主学习认识常用设计模式。	6
10	调优	掌握 Hive 性能调优的方法。	1. 了解并掌握 Hive 性能调优的方法。	由教师示范、指导，学生小组学习并练习 Hive 调优的方法。	4
11	其他文件格式和压缩方法	了解文件格式和压缩方法	1. 掌握其他人间格式和压缩方法。	由教师示范、指导学生小组学习	6
12	开发	了解 Hive 自身源代码	1. 了解 Hive 自身源代码。	由教师示范、指导学生小组学习	4
13	函数	掌握 Hive 中函数的用法	1. 掌握 Hive 中函数的用法	由教师示范、指导学生小组学习完成函数练习。	6
14	Streaming	了解 Streaming 处理数据的方式。	1. 了解 Streaming 处理数据的方式	由教师示范、指导学生自主学习	6
15	自定义 Hive 文件和记录格式	掌握自定义文件和记录格式	1. 掌握文件和记录格式	由教师示范、指导、学生自主学习	6
16	Hive 的 Thrift 服务	了解跨语言的服务开发软件框架：Thrift	1. 了解跨语言的服务开发软件框架：Thrift	由教师示范、指导学生自主学习	6
17	Hive 和亚马逊网络服务系统 (AWS)	掌握程序中各事件的综合协调处理	掌握较复杂业务的处理。	学生自主学习、小组协作	10

## 六、课程实施建议

(一) 教学建议 (从教学条件、教学方法与手段、课程资源的开发与利用、教材选用等方面进行说明)

### 1. 教学资料开发建议

- (1) 结合实际软件开发工程案例，进行课件制作、教学环境模拟。
- (2) 与企业专家合作编写符合技能培养的实验实训教材，是学生学有所用。
- (3) 提供网络资源学习平台，使学生不拘泥于固定的时间、地点学习，提高学生学习的灵活性。

### 2. 教学资源使用建议

(1) 学习包和教材，都是素材性的课程资源，但教材是知识的载体，而学习包是引导学生学习的载体。因此，要开发学习包，突出以学生为中心的学习过程，将本课程的教学内容与地区经济的发展密切结合，并体现学生职业生涯发展的需要。

(2) 教学中学生的问题、困惑、见解、情感和体验等都是动态生成的课程资源，老师应重视这些教学过程中动态生成的课程资源，并主动性和创造性地运用，使以项目为主的教学充满生机与活力。

(3) 本课程的实施，不仅应充分开发校内的课程资源，还应利用校外的课程资源，如相关企业的岗位实习和实践，尽量做到校企结合、工学结合。

(4) 积极开发教学资源：课程标准、实训指导书、授课计划等教学文件，以及课件、习题、案例库、网络方案、工具软件和网络资源等。

充分利用网络资源，了解最新的技术策略和实施方式，培养学生自主学习的能力。

### 3. 教材

选用教材：《Hive 编程指南》曹坤 人民邮电出版社

#### (二) 考核建议

考核主要由平时成绩、理论考试以及实训项目成绩组成。

##### 1. 平时成绩

平时成绩强调平时的出勤、课堂纪律、课堂表现和作业，教师通过千分制的实施可以在整个授课过程中监控学生的学习效果以及学习态度等等。平时成绩占全部考试的 30%。

##### 2. 理论考试

理论考试采用传统的笔试方式，为了更好的达到考核的目的，我们采用第三方出题的形式来组织理论考试。理论考试占全部考试的 40%。

##### 3. 项目评审

实施项目内部考核及教师考核相结合的原则，在项目实训末期，要进行项目评审，根据项目的完成情况给学生打分。项目评审成绩占全部成绩的 30%（包括学生自我评价及对同伴的评价、完成项目效果评价以及教师评价，其中自我评价（占 30%）、项目效果评价（占 30%）、教师评价（占 40%）。）

**七、需要说明的其他问题（参考资料、所需仪器、设备、教学软件等）**

1. 参考资料

(1) 《Hadoop 权威指南》 Tom White 清华大学出版社

(2) 《基于 Hadoop 与 Spark 的大数据开发实战》 人们优点出版社

2. 硬件要求：微机、网络。

3. 软件要求：JDK、Hadoop、Hive、Eclipse、VMware、XmanagerEnterprise 等软件。

## 《Spark 大数据技术与应用》课程标准

## 一、课程基本信息

课程名称	Spark 大数据技术与应用				
课程代码		学时	108	学分	6
授课时间	第四学期	适用专业	大数据应用与技术专业		
课程性质	岗位能力课程				
先修课程	Hadoop 系统基础	后续课程			

## 二、课程定位

在 park 作为下一代大数据处理引擎，在非常短的时间里崭露头角，并且以燎原之势席卷业界。Spark 对曾经引爆大数据产业革命的 Hadoop、MapReduce 的进行改进，通过学习本课程，学生能够利用 Spark 去解决一些可能会遇到但是没有办法解决的问题。能够掌握何搭建一个 Spark 集群，学会使用 Spark shell，以及编写 Spark 应用程序来解决需要并行处理的问题。

## 三、课程设计思路

本课程前两章将会带你入门，搭建基础的 Spark，了解 SPark 基本的概念。重点介绍 Spark shell。Spark shell 是开发 Spark 应用原型时非常有用的工具。后续几章则会详细介绍 Spark API、如何将 Spark 应用运行在集群上，以及 Spark 所提供的更高层的程序库支持，例如 SQL（数据库支持）和 MLlib（机器学习库）。

## 四、课程目标

## （一）能力目标

1. 增强学的大数据的理解及数据分析；
2. 培养学生对大数据生态圈的理解；
3. 培养学生发现问题、分析问题和解决问题的能力。

## （二）知识目标

1. 了解 Spark 生态圈
2. 熟练 Spark shell
3. 熟练掌握 Spark API
4. 掌握如何将 Spark 运行在集群上



5. 掌握 Scala 基本编程方法
6. 掌握 DataFrame 编程方法
7. 掌握 Spark Streaming 对 Socket、HDFS 数据进行流式处理的方法
8. 了解 Spark Streaming 与 Flume、Kafka 的整合

### (三) 素质目标

1. 培养按时完成项目的良好习惯，增强集体和社会责任感；
2. 培养团队协作和良好的沟通能力；
3. 培养学生务实、创新、诚实、守信的性格。

## 五、课程内容及要求

序号	教学内容	能力目标	知识目标	教学方法及手段	学时
1	Spark 数据分析导论	从宏观角度理解 Spark 到底是什么。	1. 了解 Spark 是什么。 2. 理解 Spark 软件栈。	引导启发学生独立完成设置工作。	9
2	Spark 下载与入门	能够下载 Spark 并在本地模式下单机运行它。	1. 掌握下载 Spark 2. 掌握 Spark 中 Python 和 Scala 的 shell。 3. 理解 Spark 核心概念简介 4. 掌握 Spark 的独立应用	由教师示范、指导，学生小组协作	9
3	RDD 编程	掌握 Spark 对数据的核心抽象 RDD	1. 掌握如何创建 PairRDD 及其转换操作 2. 掌握数据分区	由教师示范、指导、学生小组协作完成本章练习。	12
4	键值对操作	掌握如何操作键值对 RDD	理解类的继承与多态性的概念及实现	由教师示范、指导、学生小组协作完成本章练习。	9
5	数据读取与保持	掌握共享变量。	1. 了解累加器、广播变量及基于分区进行的操作。	由教师示范、指导、学生小组协作完成本章练习。	9
6	Spark 编程进阶	1. 掌握常见异常的处理 2. 学会用户自定义异常的创建	1. 理解异常的概念 2. 了解异常类的层次结构	由教师示范、指导、学生小组协作完成本章练习。	10
7	在集群上运行 Spark	掌握分布式 Spark 应用环境架构及配置	1. 理解 Spark 运行时架构 2. 掌握 spark-submit 部署应用 3. 掌握 Spark 打包代码与依赖	由教师示范、指导、学生自主学习	12
8	Spark 调优与调试	掌握如何配置 Spark 应用，了	1. 掌握如何配置 Spark 应用	由教师示范、指导、学生自主学习	9

		解在生产环境中的 Spark 工作负载如何调优和调试。	2. 了解如何调优和调试生产环境中的 Spark 工作负载。		
9	Spark SQL	掌握 SparkSQL	1. 掌握如何在应用中 使用 SparkSQL 2. 掌握用户自定义函数	由教师示范、指导、学生自主学习	10
10	Spark Streaming	掌握创建常用容器、菜单的对象、设置属性的及调用方法	1. 使用 Java 和 Scala 练习 SparkStreaming 模型。	由教师示范、指导、学生自主学习	10
11	基于 MLib 的机器学习	了解 MLib 函数库	1. 了解机器学习基础 2. 了解 MLib 中主要的算法,以及它们的输入和输出类型	由教师示范、指导、学生自主学习	12

## 六、课程实施建议

(一) 教学建议 (从教学条件、教学方法与手段、课程资源的开发与利用、教材选用等方面进行说明)

### 1. 教学资料开发建议

- (1) 结合实际软件开发工程案例, 进行课件制作、教学环境模拟。
- (2) 与企业专家合作编写符合技能培养的实验实训教材, 是学生学有所用。
- (3) 提供网络资源学习平台, 使学生不拘泥于固定的时间、地点学习, 提高学生学习的灵活性。

### 2. 教学资源使用建议

(1) 学习包和教材, 都是素材性的课程资源, 但教材是知识的载体, 而学习包是引导学生学习的载体。因此, 要开发学习包, 突出以学生为中心的学习过程, 将本课程的教学内容与地区经济的发展密切结合, 并体现学生职业生涯发展的需要。

(2) 教学中学生的问题、困惑、见解、情感和体验等都是动态生成的课程资源, 老师应重视这些教学过程中动态生成的课程资源, 并主动性和创造性地运用, 使以项目为主的教学充满生机与活力。

(3) 本课程的实施, 不仅应充分开发校内的课程资源, 还应利用校外的课程资源, 如相关企业的岗位实习和实践, 尽量做到校企结合、工学结合。

(4) 积极开发教学资源: 课程标准、实训指导书、授课计划等教学文件,

以及课件、习题、案例库、网络方案、工具软件和网络资源等。

充分利用网络资源，了解最新的技术策略和实施方法，培养学生自主学习的能力。

### 3. 教材

选用教材：《Spark 快速大数据分析》 王道远 译 人民邮电出版社

### （二）考核建议

考核主要由平时成绩、理论考试以及实训项目成绩组成。

#### 1. 平时成绩

平时成绩强调平时的出勤、课堂纪律、课堂表现和作业，教师通过千分制的实施可以在整个授课过程中监控学生的学习效果以及学习态度等等。平时成绩占全部考试的 30%。

#### 2. 理论考试

理论考试采用传统的笔试方式，为了更好的达到考核的目的，我们采用第三方出题的形式来组织理论考试。理论考试占全部考试的 40%。

#### 3. 项目评审

实施项目内部考核及教师考核相结合的原则，在项目实训末期，要进行项目评审，根据项目的完成情况给学生打分。项目评审成绩占全部成绩的 30%（包括学生自我评价及对同伴的评价、完成项目效果评价以及教师评价，其中自我评价（占 30%）、项目效果评价（占 30%）、教师评价（占 40%）。）

## 七、需要说明的其他问题（参考资料、所需仪器、设备、教学软件等）

### 1. 参考资料

- （1）《Hadoop 权威指南》 清华大学出版社
- （2）《基于 Hadoop 与 Spark 的大数据开发实战》 人民邮电出版社
- （3）《Hadoop + Spark 大数据巨量分析与机器学习整合开发实战》 林大贵 清华大学出版社

### 2. 硬件要求：微机、网络。

### 3. 软件要求：JDK、Hadoop、VMware、Eclipse 等软件。