

IOTS8.0-零基础数据分析师与大数据分析 课程

产品说明书

引领者，唯有不断超越自我

Iotek Training System 8.0

海同科技教学教研部

2019 年 3 月

一、大数据时代，数据就是生产力.....	1
二、大数据分析师，IT 职场新贵.....	2
三、大数据分析师岗位说明.....	3
3.1、工作层级：描述型—>分析型—>预测型—>规划型.....	3
3.2、职业细分方向.....	4
3.3、大数据分析师综合能力模型.....	5
四、学员对象.....	5
五、学习周期.....	5
六、学习模式.....	6
6.1、闯关+项目驱动模式、实践为主、百战成神.....	6
6.2、一周学习流程与安排.....	6
七、就业方向.....	6
八、高薪就业.....	7
九、专家讲师团队.....	9
十、课程特色.....	11
10.1、定位高端岗位.....	11
10.2、课程体系全面.....	11
10.3、项目驱动式学习，全行业大数据分析经典案例贯穿.....	11
10.4、深入大数据挖掘与分析技术体系.....	12
10.5、透析 Python 数据处理与可视化技术.....	12
10.6、行业大数据分析专家主讲，高端权威.....	14
10.7、智能学习平台+贴心学习服务，效果保证.....	14
10.8、专业【就业+猎头】服务团队，后顾之忧.....	14
十一、课程优势.....	14
11.1、专家设计岗位体系.....	14
11.2、课时设计科学合理.....	14
11.3、零基础能轻松学会.....	14
11.4、项目贯穿学习过程.....	15
11.5、8 小时/天名师视频讲解.....	15
11.6、一周三次直播互动授课.....	15
11.7、衔接企业需求定制.....	15
11.8、COT 职业导向课.....	16
11.9、全真模拟面试.....	16
十二、课程大纲.....	16
第一阶段：业务分析师（必修）.....	16
第二阶段：数据可视化（必修）.....	23
第三阶段：建模数据分析师（必修）.....	26
第四阶段：Python 数据处理与机器学习（必修）.....	32
第五阶段：大数据开发工具.....	37
第六阶段：R 数据分析与挖掘.....	40
十三、大型项目实战.....	42
项目实战 1：电商大数据运营分析.....	42
项目实战 2：连锁零售业运营数据分析.....	43

项目实战 3: 爱家房屋租赁大数据分析.....	44
项目实战 4: 智能云学习平台大数据分析.....	45
项目实战 5: 互联网分布式爬虫数据采集.....	46
项目实战 6: 互联网舆情大数据挖掘与分析.....	47
项目实战 7: 共享单车骑行大数据分析.....	48
项目实战 8: O2O 外卖点单平台大数据分析.....	48
十四、结束语.....	1

一、大数据时代，数据就是生产力

“数据是基础性、战略性资源，更是重要生产力”——国务院总理，李克强

“数据是未来竞争优势的基础”——IBM 执行总裁，罗睿兰

“数据是新的石油”——亚马逊首席科学家，安德雷斯·韦思安

“数据是这个时代最核心的资源”——阿里巴巴，马云

数据是“互联网+”时代的金矿，将成为企业 and 国家层面重要的战略资源。数据像设备、厂房、原材料一样，成为企业的重要资产，是提升企业核心竞争力、发现市场新机遇、创造新的利润增长点的重要资源。麦肯锡研究发现，充分利用大数据的企业营业额通常能够增长 60% 以上。“数据驱动决策”已成趋势。

未来最大的资源就是数据。采集、拥有数据、掌握数据挖掘、分析和利用的技术，就是掌握了“互联网+”时代的先进生产力。

大数据时代的到来，让企业深刻意识到大数据分析运用能力对企业发展的极端重要性，互联网、金融、通信、商业、制造业等领域纷纷高薪招聘大数据分析师。大数据分析师需求出现“井喷式”增长，成为 IT 新贵职业。（猎聘网）

■ 数据分析师成为大数据领域宠儿，需求占比超四成。

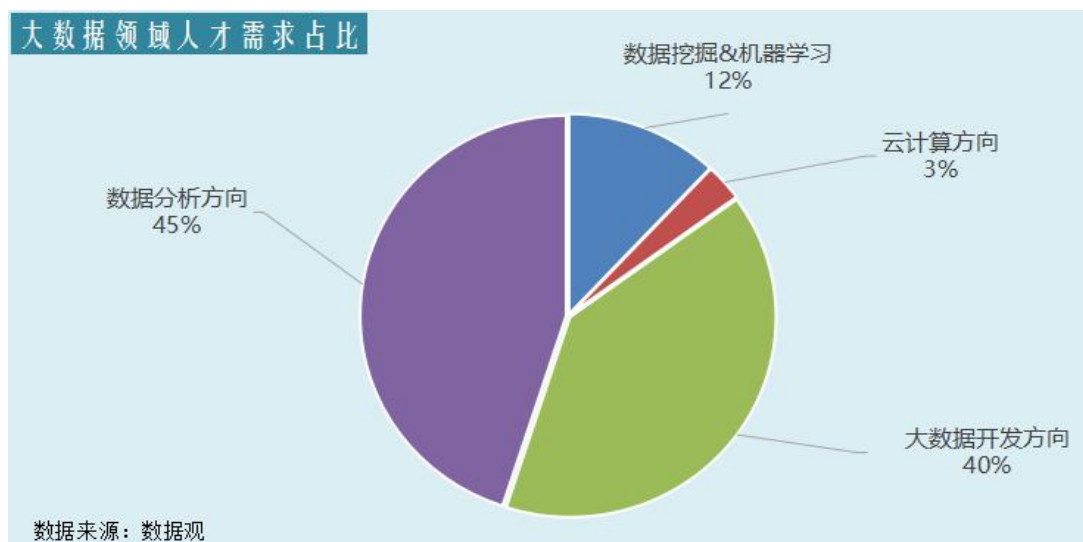


图 1、大数据领域人才需求占比 （数据来源：数据观）

二、大数据分析师，IT 职场新贵

■ 岗位关键字：高薪 + 轻编程 + 职业周期长（告别青春饭）

■ 高薪（需求旺盛、供给紧缺）

<input type="checkbox"/> 全选	<input type="button" value="申请职位"/>	<input type="button" value="收藏职位"/>	已选：大数据分析师(全文)+北京+上海		
<input type="checkbox"/> 大数据分析师/产品经理	上海纳睿宝信息技术有限公司	上海-闵行区	1.8-2.5万/月	03-29	
<input type="checkbox"/> 大数据分析师（互联网行业）	北京梅泰诺通信科技股份有限公司	北京-海淀区	1.5-3万/月	03-30	
<input type="checkbox"/> 大数据分析师	北京昊辉文化咨询有限公司	北京-朝阳区	2-2.5万/月	03-29	
<input type="checkbox"/> 大数据分析师	上海博科资讯股份有限公司	上海-普陀区	30-40万/年	03-29	
<input type="checkbox"/> 大数据分析师	北京时代凌宇科技股份有限公司	北京-朝阳区	1.5-2万/月	03-29	
<input type="checkbox"/> 大数据分析师	上海育创网络科技股份有限公司	上海-浦东新区	2-2.5万/月	03-28	
<input type="checkbox"/> 大数据分析师	北京必可测科技股份有限公司	北京	1-1.5万/月	03-28	
<input type="checkbox"/> 大数据分析师	长安新生（深圳）金融投资有限公司	北京-朝阳区	1-2万/月	03-28	
<input type="checkbox"/> 大数据分析师J10942	上海小寻科技有限公司	上海-徐汇区	1-1.5万/月	03-28	
<input type="checkbox"/> 大数据分析师	上海炫踪网络股份有限公司	上海	2.5-3万/月	03-28	
<input type="checkbox"/> 大数据分析师	创络（上海）数据科技有限公司	上海-徐汇区	2-2.5万/月	03-28	
<input type="checkbox"/> 大数据分析师	广州汇智通信技术有限公司	上海	1.5-2万/月	03-30	
<input type="checkbox"/> 法律大数据分析师	北京北大英华科技有限公司	北京-海淀区	2-2.5万/月	03-27	
<input type="checkbox"/> 大数据分析师	上海肯耐珂萨人才服务股份有限公司	上海-长宁区	30-40万/年	03-27	

图 2、大数据分析师招聘情况 （数据来源：51Job）

上海高级数据分析师 • 工资收入水平

上海高级数据分析师平均工资：¥ 19260/月，取自 883 份样本，较 2016 年，增长 30.3%



图 3、上海高级数据分析师工资水平 （数据来源：职友网）

■ 轻编程、职业周期长

传统的编程岗位，能力模型比较单一，大量的工作内容是重复性编码，工作量大且频繁加班，熬夜是常态，俗称“码农”、“IT 民工”，工作重复性大，随着年龄的增长，如果不能走向更高的架构或管理岗位，许多人都会面临职场危机——转岗、降薪、甚至被辞退，因此程序员也常被称为“吃青春饭”的职业。

与此不同的是，大数据分析师岗位对人才的综合能力要求高，包括数据分析能力、业务知识、运营知识、编程能力等。这一岗位接触面广（技术、市场、运营、管理）、视野开阔，经验非常宝贵，受年龄影响小，职业周期长。

既掌握统计知识和专业工具、又能使用 Python 等进行数据挖掘与分析、能够对企业运营决策提供科学的大数据支持，并且在工作中经常接触企业的核心业务数据，岗位重要性和持续性远高于普通编程岗位。

■ 美国国家劳工统计局数据

劳工统计局最新就业数据显示，“数据分析师”职业 2019 年预计将有 **80 万** 从业人员，**同比增长 53.4%**，大数据分析人员平均年薪达 **17.5 万** 美元，数据分析师将成为今后 5 年最热门的职业。

三、大数据分析师岗位说明

3.1、工作层级：描述型—>分析型—>预测型—>规划型

层级	说 明	工 作	角 色
描述型	描述：发生了什么？	采集数据、汇总、描述	分析员
分析型	分析：为什么发生？	研究、分析、总结	初级分析师
预测型	预测：将来会发生什么？	预测、研判、预警	中级分析师
规划型	规划：我们应该怎么做？	规划、方案、决策	高级分析师

3.2、职业细分方向

数据分析师职业可以细分为几个方向：

A、运营分析方向（偏业务）：数据分析师、数据分析主管、数据运营经理/总监

B、数据工程方向（偏技术）：数据工程师、数据挖掘工程师

C、数据产品方向（偏产品）：数据产品 PM



图 4、数据分析师职业细分方向

不同规模、不同行业的公司在架构上有所不同，但是数据分析工作职责的划分上基本都符合以上情况。

3.3、大数据分析师综合能力模型

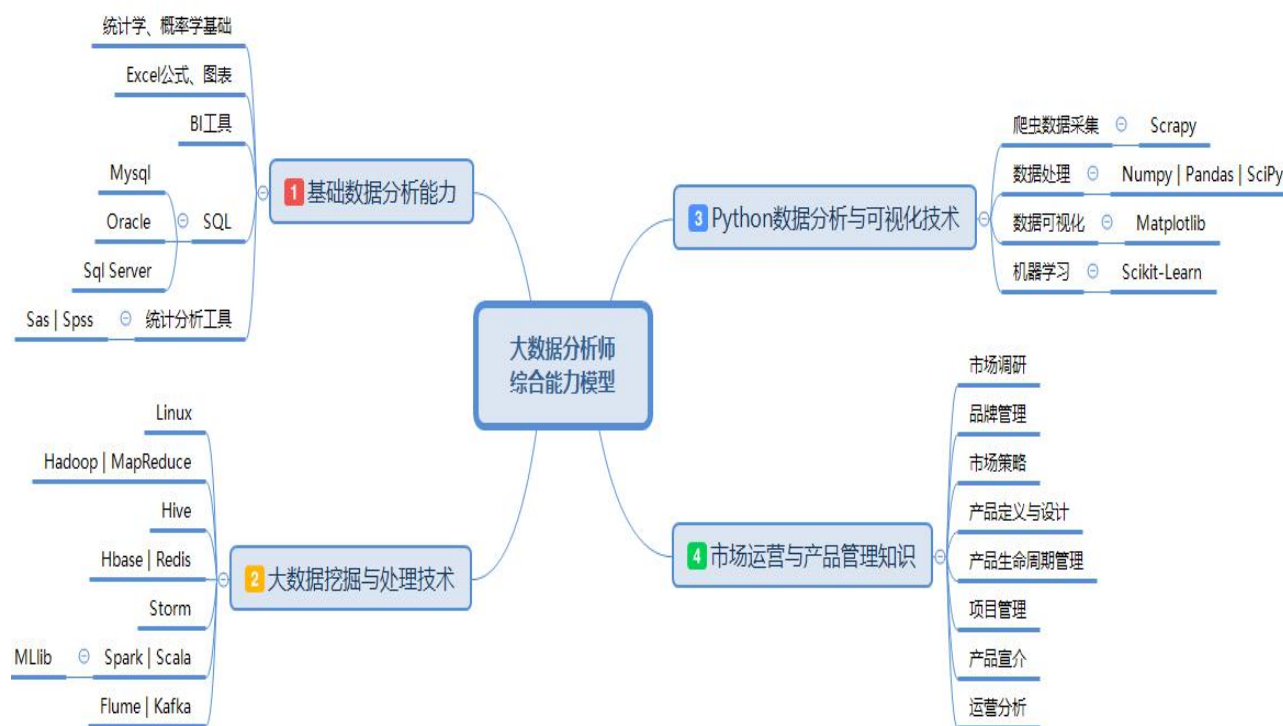


图 5、大数据分析师综合能力模型

四、学员对象

零基础，高校应届毕业生或企业在职人员，希望通过系统学习和项目实战，掌握大数据分析师技能体系，实现高薪就业或升职、跳槽。

五、学习周期

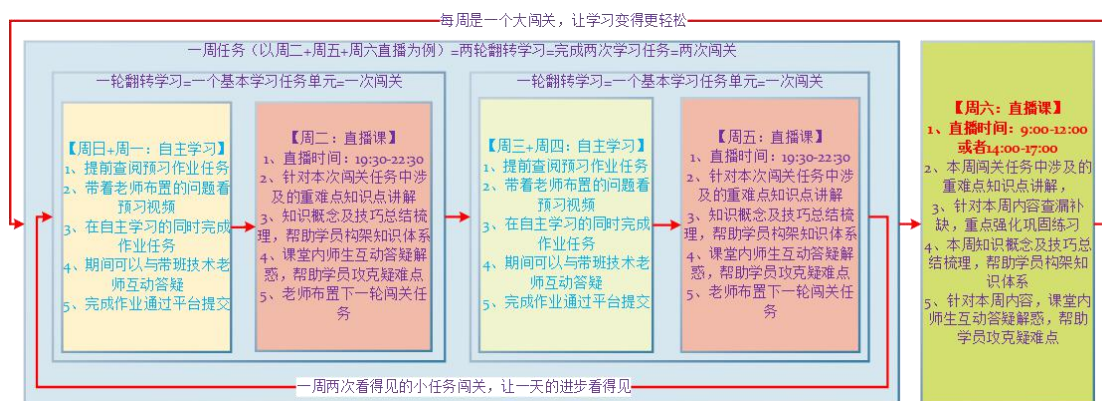
- 学习周期：5 个月【项目驱动】+【闯关】强化实战培训。
- 基于职坐标智能平台的混合式学习：平台学习+个性化学习路径+专家师资 1 对 1 辅导+海量技术资源。

六、学习模式

6.1、闯关+项目驱动模式、实践为主、百战成神



6.2、一周学习流程与安排



七、就业方向

- 大数据分析师
- 数据分析师
- BI 分析师
- 大数据工程师
- Python 工程师
- 爬虫工程师

- 数据产品经理
- 数据挖掘工程师

八、高薪就业

<input type="checkbox"/> 全选	<input checked="" type="button" value="申请职位"/>	<input checked="" type="button" value="收藏职位"/>	已选：大数据分析师(全文)+北京+上海		
<input type="checkbox"/> 大数据分析师/产品经理	上海纳睿宝信息技术有限公司	上海-闵行区	1.8-2.5万/月	03-29	
<input type="checkbox"/> 大数据分析师(互联网行业)	北京梅泰诺通信技术股份有限公司	北京-海淀区	1.5-3万/月	03-30	
<input type="checkbox"/> 大数据分析师	北京昊辉文化咨询有限公司	北京-朝阳区	2-2.5万/月	03-29	
<input type="checkbox"/> 大数据分析师	上海博科资讯股份有限公司	上海-普陀区	30-40万/年	03-29	
<input type="checkbox"/> 大数据分析师	北京时代凌宇科技股份有限公司	北京-朝阳区	1.5-2万/月	03-29	
<input type="checkbox"/> 大数据分析师	上海育创网络科技股份有限公司	上海-浦东新区	2-2.5万/月	03-28	
<input type="checkbox"/> 大数据分析师	北京必可测科技股份有限公司	北京	1-1.5万/月	03-28	
<input type="checkbox"/> 大数据分析师	长安新生(深圳)金融投资有限公司	北京-朝阳区	1-2万/月	03-28	
<input type="checkbox"/> 大数据分析师J10942	上海小寻科技有限公司	上海-徐汇区	1-1.5万/月	03-28	
<input type="checkbox"/> 大数据分析师	上海炫踪网络股份有限公司	上海	2.5-3万/月	03-28	
<input type="checkbox"/> 大数据分析师	创络(上海)数据科技有限公司	上海-徐汇区	2-2.5万/月	03-28	
<input type="checkbox"/> 大数据分析师	广州汇智通信技术有限公司	上海	1.5-2万/月	03-30	
<input type="checkbox"/> 法律大数据分析师	北京北大英华科技有限公司	北京-海淀区	2-2.5万/月	03-27	
<input type="checkbox"/> 大数据分析师	上海肯耐珂萨人才服务股份有限公司	上海-长宁区	30-40万/年	03-27	

数据来源：51Job

北京高级数据分析师 • 工资收入水平

北京高级数据分析师平均工资：¥ 24090/月，取自 1355 份样本，较 2016 年，增长 35.4%



上海高级数据分析师 • 工资收入水平

上海高级数据分析师平均工资：¥19260/月，取自 883 份样本，较 2016 年，增长 30.3%

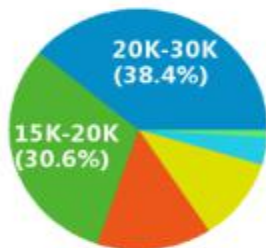
最新招聘

99+

工资收入

就业形势

岗位职责



区间	占比
4.5K-6K	0.1%
6K-8K	1%
8K-10K	3.5%
10K-15K	14.4%
15K-20K	30.6%
20K-30K	38.4%
30K-50K	11%

¥ 19260

近1年 1883 份样本 / 可信度：高

你觉得该统计数据准确吗？

偏低

8票 (57%)

靠谱

4票 (29%)

偏高

2票 (14%)

上海爬虫工程师 • 工资收入水平

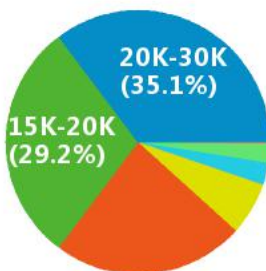
上海爬虫工程师平均工资：¥17540/月，取自 1180 份样本

最新招聘

99+

工资收入

就业形势



区间	占比
3K-4.5K	0.4%
6K-8K	2.2%
8K-10K	2.5%
10K-15K	23.2%
15K-20K	29.2%
20K-30K	35.1%
30K-50K	6.6%

¥ 17540

近1年 1180 份样本 / 可信度：高

你觉得该统计数据准确吗？

偏低

9票 (64%)

靠谱

2票 (14%)

偏高

3票 (21%)

相似岗位工资收入

数据挖掘工程师(上海) ¥22220

技术支持(上海) ¥7880

保险代理人(上海) ¥12150

销售顾问(上海) ¥10440

平面设计师(上海) ¥7840

投资顾问(上海) ¥10430

技术支持工程师(上海) ¥8070

高级数据挖掘工程师(上海) ¥28550

电商运营经理(上海) ¥14480

运营专员(上海) ¥7350

以人民币计，截至2019-02-22

¥4000

¥14000

¥24000

图 7、大数据分析师薪资收入情况 （数据来源：职友网）

九、专家讲师团队

曹宇苗 (Amily)

资深数据分析师，10 年金融/互联网商业数据分析经验，国家统计局认证高级市场调查分析师，教育部《知识经济与可持续发展》、《可持续发展核算》课题研究、撰稿人，东北财经大学统计学硕士，著有《tableau 商业数据分析实战教程》并参与主讲网易-中国大学 Mooc《SPSS 金融大数据分析》课程。

工作经历：

- 上海大钱信息技术有限公司 数据分析经理
- Sudler MDS (Shanghai) Co., Ltd 数据分析师
- 上海丽人丽妆有限公司 数据分析师
- 上海浦东发展银行 风控分析师

参与的主要项目：

- 日本大型餐饮网站 GUDUMAMI 上海版市场数据分析
- 大钱包 APP 数据分析指标建立
- 大钱包 APP 用户画像分析
- 上海浦东发展银行贷款风控模型
- 上海大众互联网广告投放数据模型分析与优化报告
- 教育部《知识经济与可持续发展》、《可持续发展核算》项目课题



朱运顺 (Eric)

- 知名互联网公司架构师
- 精通 Java、Python 语言
- 精通 springcloud 分布式系统微服务架构
- 曾设计“胖猫宝”基金平台后台架构
- 毕业于合肥工业大学软件工程专业
- 8 年以上软件开发经验

参与项目：

- 泛微 OA 项目（泛微网络）
- “胖猫宝”基金项目（胖猫宝）
- “大昌共享”共享汽车（大昌行集团）



范佳菁（Edmund）

数据分析师教学 5 年，有金融/互联网商业数据分析经验，精通 excel，ppt，tableau，python 等工具，精通 Access，Linux，Mysql 系统，对日常经营和平台数据进行采集和加工，通过数据洞察业务经营状况和产品使用情况，个人专注于数据挖掘、大数据分析技术。教学风格风趣幽默，广受学员好评。

工作经历：

- 上海交通银行卡中心
- 上海沸蓝信息科技有限公司
- 上海海同信息科技有限公司
- 参与的主要项目：
- 信用卡客户价值分析
- 用户消费行为分析
- 银行贷款风控分析
- 用户消费需求分析
- 银行信用卡欺诈与拖欠行为和客户分析
- 联合办公市场现状分析



十、课程特色

10.1、定位高端岗位

面向互联网、金融、电子、通信、高端制造等行业需求，培养紧缺的**高端大数据分析人才**而非**初级数据分析人员**。

10.2、课程体系全面

IOTS8.0 大数据分析师高端课程，由业内顶级大数据分析专家设计并参与授课，打造最高端的大数据分析师培养体系。



10.3、项目驱动式学习，全行业大数据分析经典案例贯穿

不同于许多机构的数据分析师课程偏重理论学习、内容枯燥，IOTS8.0 大数据分析师高端课程围绕**四十九个经典大数据分析项目案例**展开学习，涵盖金融、

互联网、电信、制造、教育、商业、房地产、媒体娱乐等多个行业，直接面对未来工作需求，告别“纸上谈兵”、真正实现“学习=工作经验积累”，学习效果远高于传统培训课程。

10.4、深入大数据挖掘与分析技术体系

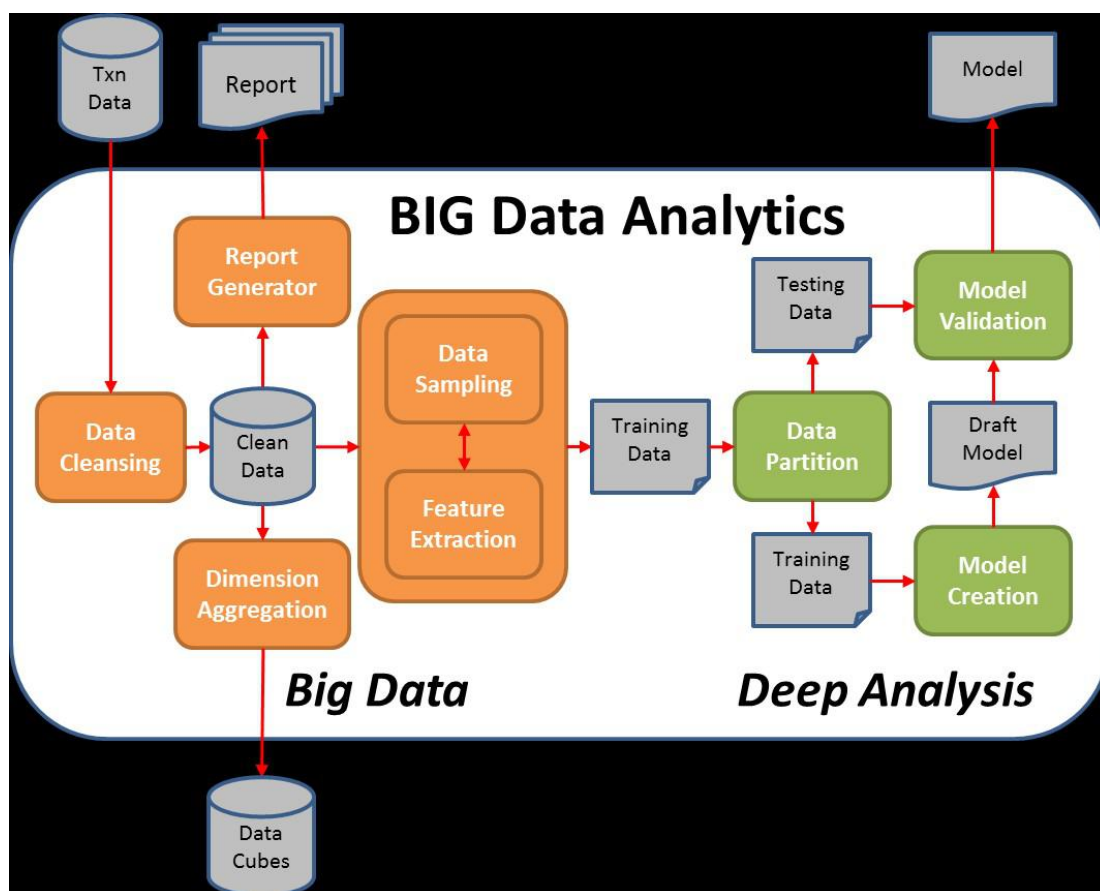


图 9、大数据分析流程

10.5、透析 Python 数据处理与可视化技术

- Python 是高级数据分析人员的必备技能，Python 在数据处理与可视化方面有非常丰富的库，功能强大。
- 我们的课程涵盖——

- ✓ Python 数据可视化
- ✓ Python 爬虫数据采集
- ✓ Python 机器学习
- ✓ Python 数据处理

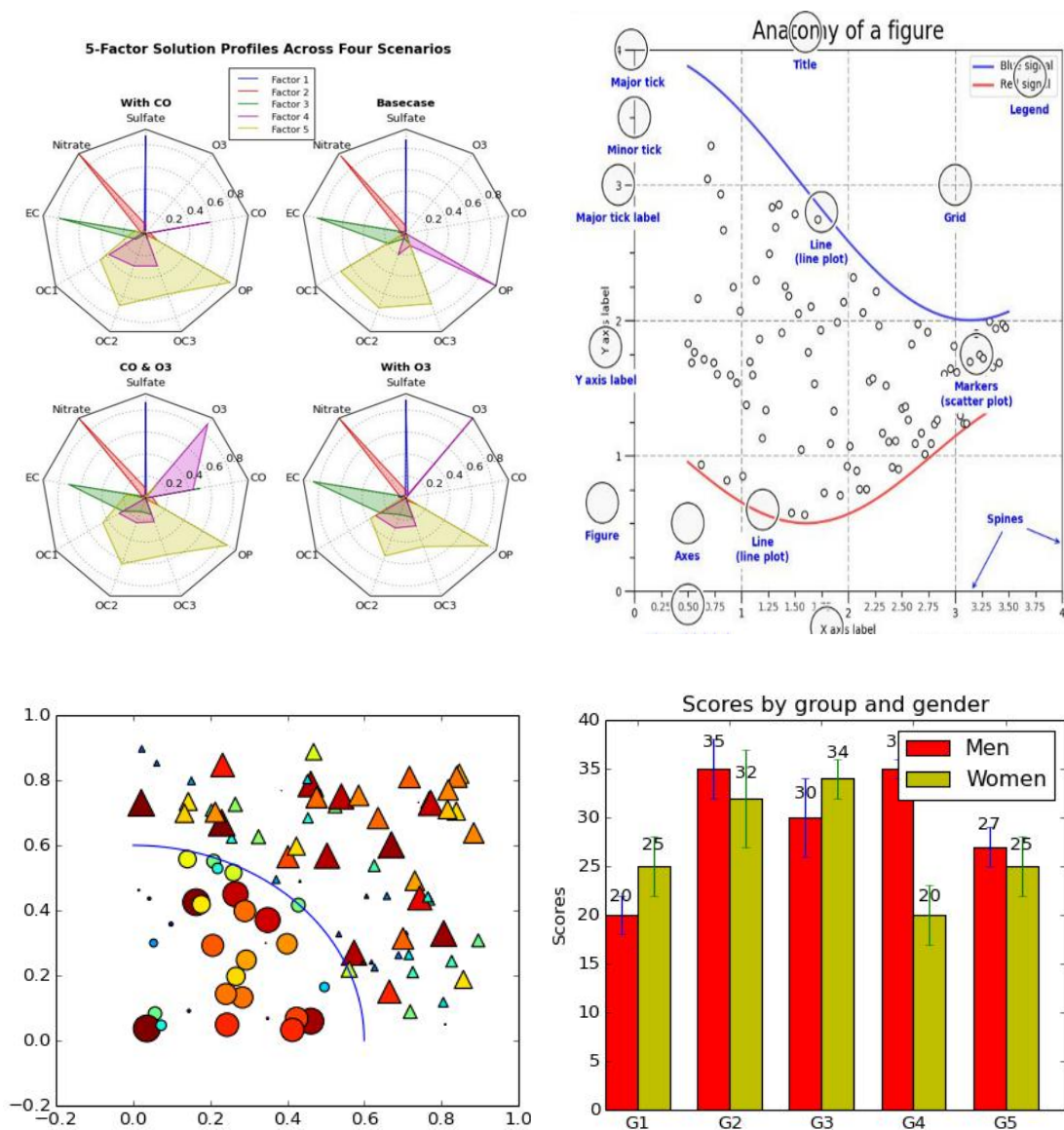


图 10、Python 数据可视化、数据分析案例图

10.6、行业大数据分析专家主讲，高端权威

10.7、智能学习平台+贴心学习服务，效果保证

10.8、专业【就业+猎头】服务团队，后顾之忧

十一、课程优势

11.1、专家设计岗位体系

数据分析师与大数据分析课程视频内容由职坐标教研专家联合多位国内知名 IT 企业技术总监根据行业岗位需求共同定制，教学内容 符合市场对数据分析师与大数据分析工程师的需求；基于前端+后端+ 设计+产品的复合型开发人才需求，职坐标培养企业需求量大、就业率高、精通数据分析师与大数据分析的中高级工程师，由浅入深 逐层讲授当前主流的数据分析师与大数据分析技术。

11.2、课时设计科学合理

数据分析师与大数据分析课程设计符合人体记忆和知识消化曲线；课时安排科学合理，在有限的时间内充分吸收专业知识；零基础的学员能轻松入门学习数据分析师与大数据分析的系统知识，金牌讲师直播精讲，核心知识点重点讲解；有基础的学员可以巩固加强自身的编程能力，通过项目操作提升岗位技能；经过系统化 4-6 个月的学习，掌握数据分析师与大数据分析的核心技术要点。

11.3、零基础能轻松学会

通过 9 年教学实践经验累积，职坐标自主研发了数据分析师与大数据分析工程师培训数字课程；全套课程内容分阶段分易难点，由浅入深、从简至难、循序渐进；优质的课程内容、清晰的视频讲解、把行业术语化繁为简，确保学员都能听得懂；针对编程基础相对偏弱的学员，职坐标将安排老师一对一辅导，助其克

服困难，针对性解决技术难题、矫正学习方法、增强学习兴趣、提高学习质量。

11.4、项目贯穿学习过程

数据分析师与大数据分析项目实战演练，让你学有所成学有所用；通过项目实操系统归纳章节内容，360° 排查知识盲点，弥补技能弱项；在项目操作的过程中，项目老师全程跟踪进度，通过项目问答与项目功能演示，反映项目实战能力。项目指导老师会在第一时间进行指正分析，提高学员专业技能。

11.5、8 小时/天名师视频讲解

授课团队由 IBM、微软、华为等知名企业软件专家组成，拥有 10 年以上开发及项目管理经验。顶级技术授课名师的开发经验合计达 52 年，授课时长合计 52000 小时，累计培训学员近万人，企业软件专家将编程思维及企业应用实践技能毫无保留的传授于你。讲师单课日均 2 小时视频课程讲解，全面解决你的学习疑难。

11.6、一周三次直播互动授课

每年 1000 多场在线直播课程，每月 100+课时名师直播互动授课，职坐标坚持每周不低于三次的课程直播互动，让学员能够在直播过程中及时与讲师交流解决学习中的疑难杂症，更有企业技术主管与学员的双向互动，无论是身处何处都能与名企软件专家零距离交流互动。

11.7、衔接企业需求定制

职坐标和全国近万家优质互联网+IT 企业建立了紧密的长期合作，结合企业需求定制数据分析师与大数据分析人才培养计划，教学结果以高薪就业为导向；资深教学导师一对一直接教学及项目指导，结合 IT 学友互动、企业技术大咖经验分享，让学员充分感受企业对于数据分析师与大数据分析工程师的需求，让学习更有趣更有效；职坐标已成功输送了数万高技能的实用型专业数据分析师与大数据分析人才。

11.8、COT 职业导向课

在面试技巧、心态建设、简历包装、技能补充等方面给予学员全面的指导，增强自信、提高面试成功率；职坐标 COT 职业导向课的授课老师均是企业集团从业多年的人事 HR，有着非常丰富的培训经验，由面至点，直击学员薄弱环节，对学员综合职业素质进行的训练，对学员步入企业，在企业生存发展有巨大帮助。

11.9、全真模拟面试

通过职坐标数据分析师与大数据分析岗位测评后，IT 企业技术主管、名企 HR 全真模拟企业面试流程，一对一技能与软技能的考查与辅导；职坐标针对学员在面试时可能遇到的问题，全真模拟企业面试流程进行面试，找出学员在面试时存在的问题，给出建议。学员再配合强化改善，对症下药，真正提高面试通过率。

十二、课程大纲

第一阶段：业务分析师（必修）

第一阶段：业务分析师
1、Excel 数据分析
第一章 Excel 基础 01-数据分析前言 02-Excel2016 初识 03-数据快速录入

04-自动填充

05-区域选择

06-复制粘贴

07-数据格式

08-查找替换

09-拆分和冻结

10-隐藏和分组

11-合并计算

12-分列与合并

13-排序和筛选

14-分类汇总

15-定位

16-导入数据

17-插入批注

18-Excel 表结构

19-条件格式

20-隐藏复制粘贴

第二章 Excel 函数

01-Excel 函数

02-绝对引用和相对引用

03-99 乘法表

04-逻辑函数-逻辑值

05-逻辑函数-and or not

06-逻辑函数-if 函数

07-逻辑函数-if 函数嵌套

08-逻辑函数-个人所得税

09-文本函数-len

10-文本函数-find 和 search

11-文本函数-left.right 和 mid

12-文本函数-trim 和 clean

13-文本函数-综合练习一

14-文本函数-concatenate

15-文本函数-text

16-文本函数-综合练习二

17-时间日期函数-本质

18-时间日期函数-提取函数

19-时间日期函数-weekday

20-时间日期函数-datedif

21-统计函数-count 系列

22-统计函数-sum 系列

23-统计函数-average 系列

24-统计函数-通配符

25-统计函数-练习

26-统计函数-练习

27-统计函数-练习

28-统计函数-练习

29-统计函数-练习

30-统计函数-median 和 mode

31-统计函数-rank

32-统计函数-最大、最小

33-查找函数-vlookup 精确查找

34-查找函数-vlookup 模糊查找

35-查找函数-hlookup

36-查找函数-vlookup 练习

37-查找函数-match 和 index

38-查找函数综合练习

39-引用函数-indirect

40-数学函数-round

41-数学函数-mod

42-数学函数-rand 系列

43-数学函数-convert

44-信息函数-is 系列

45-信息函数-iferror

46-数组 01、数据分析概述—数据分析流程、方法、在企业管理中的应用

2、Excel 数据分析高级

第一章 数据透视表

01-数据透视表简介

02-经典数据透视表布局

03-数据透视表的项

04-父字段和子字段

05-注意事项

06-数据透视表工具简介

07-数据透视表选项卡

08-数据透视表-日期设置

09-数据透视表-自定义分组

10-数据透视表-计算字段

11-数据透视表-计算项

12-数据透视表-字段设置

13-数据透视表-切片器和日程表

14-数据透视图

第二章 Excel 画图

01-Excel 作图-柱形图

02-Excel 作图-饼图

03-Excel 作图-三维簇状柱形图

04-Excel 作图-迷你图

05-Excel 作图-销量对比图

06-Excel 作图-三维饼图

07-Excel 作图-面积图

08-Excel 作图-百分比堆积柱状图

09-Excel 作图-条形图

10-Excel 作图-组合图

3、MySQL 简介及安装配置

第 1 章 数据库系统概论

01-序言

02-数据管理技术的产生和发展

03-关系型数据库中常用概念

第 2 章 MySQL 简介及安装配置

01-序言

02-MySQL 简介

03-MySQL 下载与安装

04-登录 MySQL

05-数据库管理

第 3 章 MySQL 数据表管理(上)

01-序言

02-MySQL 数据类型

03-创建数据表

04-为列添加约束

第 4 章 MySQL 数据表管理(下)

01-序言

02-查看数据表

03-修改数据表

04-删除数据表

第 5 章 数据管理

01-序言

02-插入数据

03-修改数据

04-删除数据

第 6 章 数据查询（上）

01-序言

02-查询基础

03-简单查询

04-条件查询

05-对查询结果排序

第 7 章 数据查询（下）

01-序言

02-汇总和分组数据

03-连接查询

04-子查询

05-联合查询

第 8 章 MySQL 函数

01-序言

02-日期与时间函数

03-数学函数

04-字符串函数

05-系统函数

第二阶段：数据可视化（必修）

第二阶段：数据可视化

1、Tableau 商业智能分析

第 1 章 基础操作

01-tableau 简介

02-tableau 操作-排序

03-tableau 操作-分层

04-tableau 操作-分组

05-tableau 操作-参数设置

06-tableau 操作-聚合函数 sum

07-tableau 操作-count 函数

08-tableau 操作-日期函数

09-tableau 操作-逻辑函数

10-tableau 操作-函数中嵌入参数

第 2 章 可视化

01-tableau 操作-地图

02-tableau 操作-条形图

03-tableau 操作-线形图

04-tableau 操作-饼图

05-tableau 操作-复合图

06-tableau 操作-嵌套条形图

07-tableau 操作-动态图

08-tableau 操作-热图

09-tableau 操作-突显图

10-tableau 操作-散点图

11-tableau 操作-甘特图

12-tableau 操作-标靶图

13-tableau 操作-箱型图

14-tableau 操作-瀑布图

15-tableau 操作-直方图

16-tableau 操作-帕累托图

17-tableau 操作-填充气泡图

18-tableau 操作-文字云

19-tableau 操作-树状图

20-新建一个仪表盘

第 3 章 案例

01-学校教育水平评估

02-城市教育水平评估

03-估值分析动态仪表板

04-职工年龄分析

05-网站客户分析

06-创建一个故事

2、Power BI 数据可视化

第一章、power BI 数据处理

1. Power BI 产品介绍
2. Power BI 产品前景
3. 各种不同类型的数据获取
4. power BI 编辑查询
5. power BI 合并表
6. power BI 分类汇总表
7. power BI 长表宽表互换
8. power BI 上传数据集
9. 编辑数据集之间的关系
10. 建立数据表之间的关系
11. 新建度量值
12. 新建列
13. 新建数据
14. 新建表

15. 筛选器方向

16. 建立一对一关系

第二章、power BI 生成可视化报表

1. 卡片图

2. 条形图

3. 组合图

4. 散点图

5. 子弹图

6. 马表图

7. 地图可视化

8. 漏斗图

9. 跑马图

10. 脉动图

11. 文字云

第三章、power service

1. 显示、收藏夹和特色仪表板

2. 分享给同事

3. 多人协作

第三阶段：建模数据分析师（必修）

第三阶段：建模数据分析师

1、统计学分析理论

第一章 统计学及统计数据

01-统计学概述

02-统计数据的类型

03-统计中的基本概念

第二章 数据的收集

01-数据的抽样

02-抽样的组织形式

第三章 数据的图表展示

01-数据的图表展示

第四章 数据的概括性度量

01-数据的概括性度量

02-数据的集中趋势

03-数据的离散趋势

04-偏态和峰态

第五章 概率与概率分布

01-随机事件和样本空间

02-概率的定义

03-条件概率

04-乘法公式

05-全概率公式

06-独立性

07-贝叶斯公式

08-离散型随机变量及其分布律

09-连续型随机变量及其分布律

第六章 统计量及其抽样分布

01-统计量

02-其他常见分布

03-样本均值的分布与中心极限定理

04-其他抽样分布

第七章 参数估计

01-估计量和估计值

02-点估计

03-区间估计

04-评价估计量的标准

05-一个总体参数的区间估计

06-两个总体参数的区间估计

07-样本量的确定

第八章 假设检验

01-假设检验基本问题

02-总体参数的检验

2、spss 数据分析

第 1 章 软件安装

01-软件安装

第 2 章 spss 简介

01-SPSS 简介

第 3 章 spss 基础操作

01-SPSS 基础操作 1

02-SPSS 基础操作 2

03-SPSS 基础操作 3

04-SPSS 基础操作 4

第 4 章 T 检验

01-单样本 t 检验

02-两样本 t 检验

03-配对 t 检验

第 5 章 方差分析

01-方差分析

02-单因素方差分析

03-双因素方差分析

第 6 章 卡方检验

01-卡方检验

第 7 章 相关分析

01-相关分析

第 8 章 回归分析

01-一元线性回归

02-多元线性回归

03-曲线回归

第 9 章 logistics 回归

01-logistics 回归分析

第 10 章 时间序列

01-时间序列分析 1

02-时间序列分析 2

第 11 章 主成分与因子分析

01-主成分与因子分析

第 12 章 聚类分析

01-聚类分析

第 13 章 RFM 模型

01-RFM 模型

3、SPSS Modeler 数据挖掘

第 1 章 基础部分

01-数据挖掘及 modeler 简介

02-modeler 软件界面介绍

03-执行流和源节点

04-缓存节点和超节点

05-文件集成

06-缺失值处理

07-字段排序

08-分类汇总

09-变量变化

10-变量派生

第 2 章 模型篇

01-KNN 算法原理介绍

02-KNN 案例 1

03-KNN 案例 2

04-朴素贝叶斯算法

05-贝叶斯网络概述

06-数据挖掘体系介绍

07-线性回归

08-logistics 回归

09-决策树入门

10-理解信息熵

11-ID3 算法

12-C4.5 算法

13-cart 树算法介绍

14-chart 树算法介绍

15-回归树和模型树介绍

16-BP 神经网络

17-BP 神经网络算法演示

18-BP 神经网络演示

19-SVM 原理介绍

20-聚类分析

21-聚类分析案例

22-apriori 算法

23-关联规则案例

24-时间序列案例 1

25-时间序列案例 2

第四阶段：Python 数据处理与机器学习（必修）

第四阶段：Python 数据处理与机器学习

1、Python 语言基础

01-python 入门

02-写一个 python 程序

03-基础语法之数据类型

04-基础语法之变量和常量

05-基础语法之运算符初级

06-基础语法之运算符高级

07-条件判断

08-字符串编码和格式化

09-字符串

10-list 列表

11-tuple 元组

12-dict 字典

13-set 集合

14-循环

15-函数及调用

16-函数参数详解 1

17-函数参数详解 2

2、Python 高级特性

01-切片

02-迭代

03-推导式

04-迭代器

05-生成器

06-文件 IO

07-文件和目录

08-StringIO 和 BytesIO

09-序列化

10-类和实例

11-实例详解

12-继承，多态和操作对象函数

13-__slots__和@property

14-定制类

15-元类和多重继承

16-自定义模块及搜索路径

17-内建模块 collections

18-内建模块 base64 和 time

19-其它内建模块简要介绍

20-第三方模块安装和使用

21-虚拟环境 virtualenv

22-第三方常用模块简要介绍

3、Python 爬虫开发

01-从一个项目开始学习网络爬虫

02-详解 http 协议

03-urllib2 入门

04-urllib 和 urllib2

05-urllib2 高级用法

06-urlError 处理

07-构建 cookie

08-正则表达式

09-用面向对象的方法抓取糗事百科

10-requests 模块常用功能

11-requests 模块高级功能

12-beautiful soup 基本功能

13-beautiful soup 高级功能

14-从登录到抓取人人网图片

15-PhantomJS

16-Selenium

17-模拟登陆并抓取知乎问题信息

18-Scrapy 爬虫框架

19-案例-Scrapy 抓取豆瓣热门电影信息

4、Python 数据分析

第一章 Numpy

01-numpy: Numpy 介绍

02-numpy: 多维数组

03-numpy: 数据类型设定与转换

04-numpy: 索引与切片

05-numpy: 多维数组索引、切片

06-numpy: 数组文件输入输出

07-numpy: 获取数组的属性

08-numpy: 合并数组

09-numpy: 基本的矩阵运算

10-numpy: Array Creation Routines

11-numpy: 从已知数据中构造数组

第二章 Pandas

01-Pandas: 字典

02-Pandas: ndarray

03-Pandas: Series 字典

04-Pandas: ndarrays 或 lists 字典

05-Pandas: 带字典的列表

06-Pandas: DataFrame.from_ 方法

07-Pandas: 列选择, 添加, 删除

08-Pandas: 面板数据

09-Pandas: Panel 构成

10-Pandas: Panel 的未来

第三章 Matplotlib

01-Matplotlib: 简介

02-Matplotlib: 图例、标题和标签

03-Matplotlib: 条形图和直方图

04-Matplotlib: 散点图

05-Matplotlib: 堆叠图

06-Matplotlib: 饼图

07-Matplotlib: 从文件加载数据

08-Matplotlib: 从网络加载数据

09-Matplotlib: 时间戳的转换

10-Matplotlib: 基本的自定义

11-Matplotlib: 颜色和填充

12-综合案例: 股票数据统计图表

第五阶段：大数据开发工具

第五阶段：大数据开发工具

课程 1、Hadoop 大数据平台应用

第 1 章 Hadoop

01-hadoop 系统功能和结构功能介绍

02-secureCRT 和 FileZilla 的使用介绍

03-hadoop 集群安装过程

04-hadoop 配置总结

05-hadoop 集群数据观察

06-向 hadoop 集群提交任务演示

第 2 章 HDFS

01-hdfs 指令实践

02-Java 操作 hdfs(一)

03-Java 操作 hdfs(二)

04-Java 操作 hdfs(三)

第 3 章 MapReduce

01-mapreduce 编程框架

02-wordcount 解决思路分析

03-wordcount 代码编写（上）

04-wordcount 代码编写（下）

05-wordcount 集群运行和观察

06-单词去重代码

课程 2、大数据仓库应用

第 1 章 Hive

01-数据仓库

02-hive 框架

03-hive 和 hadoop 之间的关系

04-hive 数据存储

05-安装 mysql

06-安装 hive

07-hive 初体验

08-mysql 中查看元数据信息

09-ddl (1)

10-ddl (2)

11-ddl (3)

12-ddl (4)

13-dml

14-join 操作

15-hive 常用函数

16-hive 实践_Java 代码操作 hive

第 2 章 Hbase

01-hbase 介绍_概念

02-hbase 介绍-表结构

03-hbase 系统结构和存储

04-zookeeper 介绍

05-hbase 的安装和配置

06-hbase shell 指令介绍

07-shell 命令实践

课程 3、大数据平台 Spark 工具

第 1 章 Scala

01-scala 介绍及环境安装

02-数值类型

03-if 和 match

04-块表达式和循环

05-方法和函数（一）

06-方法和函数（二）

07- 数组

08- list

09- map

10-Set

11-元组

12-wordcount 练习

第 2 章 Spark

01-spark 介绍和集群搭建

02-intellij 上编写程序并提交集群

03-spark shell 编写 wordcount

04- rdd 概念介绍

05-RDD 和算子

06-rdd 算子简单应用（一）

07-rdd 算子简单应用（二）

08-高级算子（一）

09-高级算子（二）

10-判断位置_背景讲述

11-统计家庭住址和工作地点（一）

12-统计家庭住址和工作地点（二）

第六阶段：R 数据分析与挖掘

第六阶段：R 数据分析与挖掘

1、R 基础

第一章：R 语言的入门简介

- 1, 课程体系的介绍
- 2, R 语言的入门简介
- 3, R 语言的入门简介-R 包
- 4, R 语言的入门简介-帮助

第二章：R 语言基础

- 1, 向量
- 2, 函数
- 3, function 函数
- 4, 逻辑、布尔运算符
- 5, 因子
- 6, 日期函数
- 7, 矩阵
- 8, 数组

- 9, 列表
- 10, 数据框
- 11, 数据框相关函数
- 12, 数据的导入导出
- 13, 控制流
- 14, 高级循环函数
- 15, 数据处理函数
- 16, 数据处理包-dplyr 包

2、R 作图

第三章 R 基础作图

- 1.R 作图—yans
- 2.R 作图—文字元素
- 3.R 作图—点、线
- 4.标题、图例、网格线
- 5.常见图

第四章 ggplot 作图

- 1.数据准备
- 2.qplot 函数
- 3.ggplot 函数

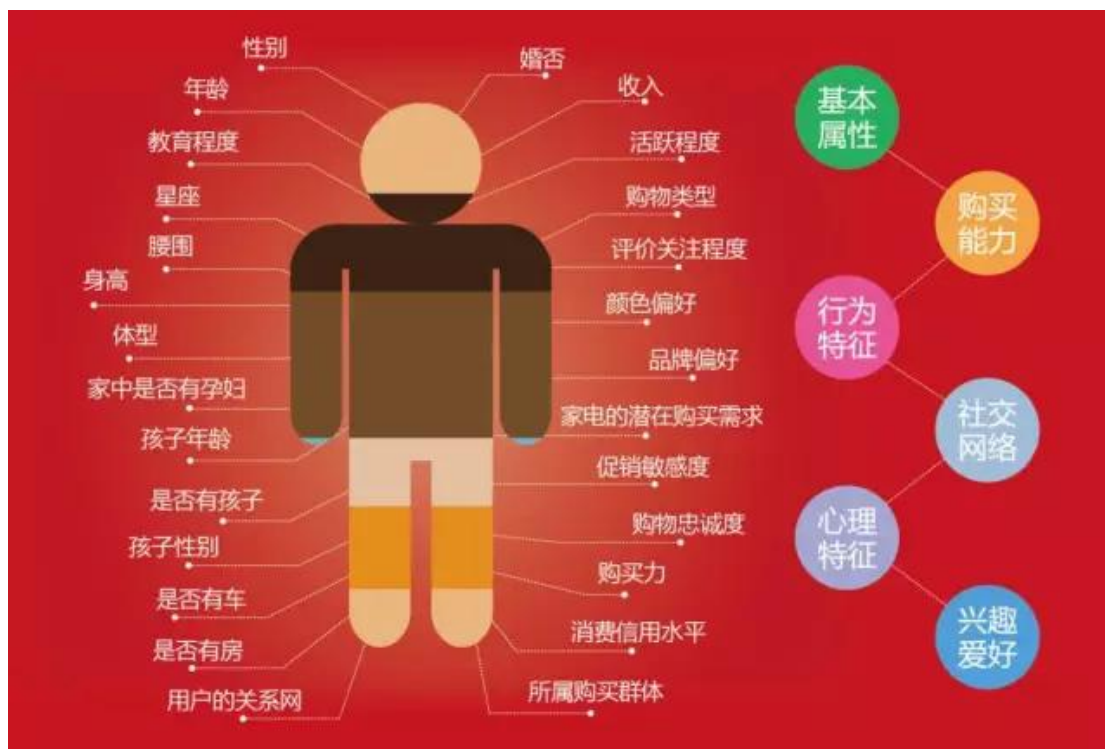
3、R 数据挖掘

第五章 R 模型

- 1.犯罪率的预测——线性回归
- 2.股票预测-logistics
- 3.车险购买预测-KNN
- 4.贷款风险-C5.0
- 5.kmeans 聚类
- 6.层次聚类
- 7.购物车案例-关联规则

十三、大型项目实战

项目实战 1：电商大数据运营分析



购物街服装客户画像

购物街服装客户三大特点



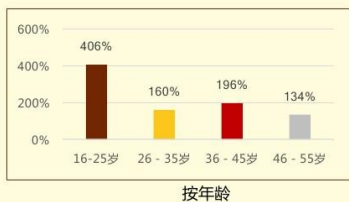
+



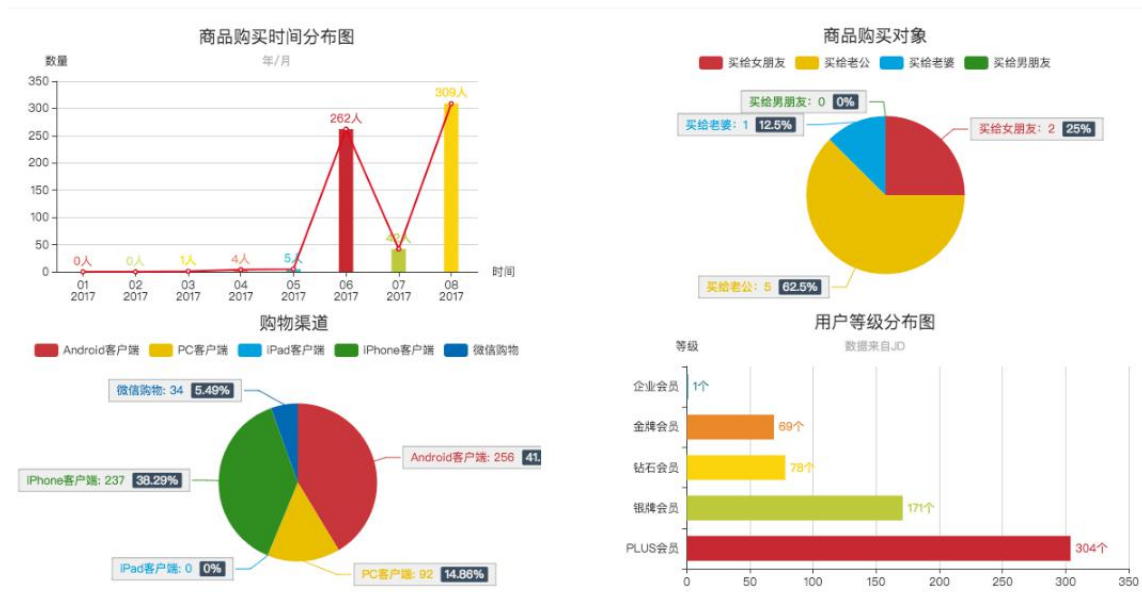
+



2017年相比2015年用户数增长速度



评论分析



项目实战 2：连锁零售业运营数据分析

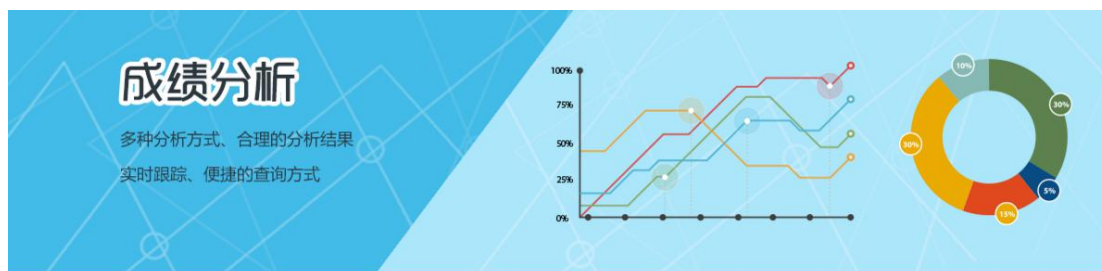




项目实战 3：爱家房屋租赁大数据分析



项目实战 4：智能云学习平台大数据分析



系统设置
对统计分析参数、试卷分析参数进行默认设置，设置文理班级。

新建考试
分权限新建考试，填写试卷属性，查看填写状态等。

成绩录入
在线录入、模板导入、阅卷系统同步等三种方式录入成绩，随时查看录入状态。

成绩查询
分权限查询学生历次考试的原分数，方便导出下载。

统计分析
多种统计方式并存，满足学校各种统计分析报表要求，各种分析结果都有图、表两种方式展现。

学生成绩单
方便导出、打印学生因自主招生或出国留学所需的正规成绩单。

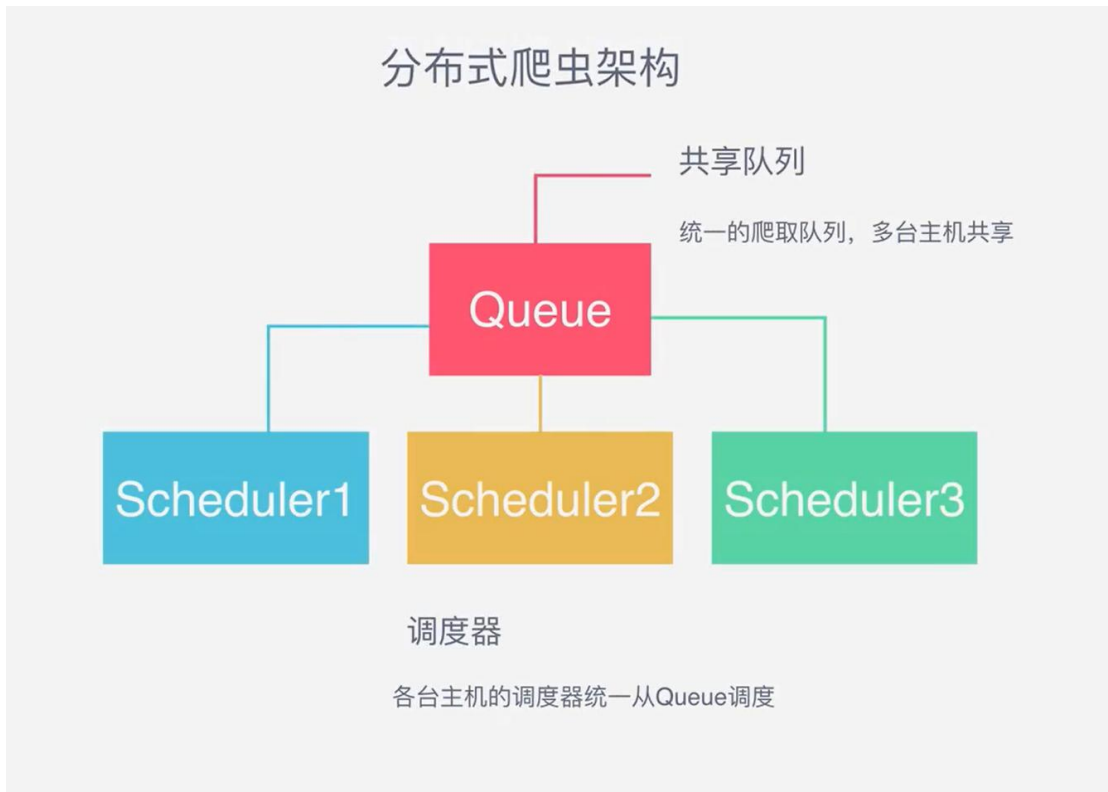
成绩跟踪
跟踪、对比历次考试结果，形象直观的展示年级、班级、个人的成绩变化。

试卷分析
通过难度、区分度等一些参数衡量试卷出质量。

在线咨询
ONLINE CONSULTATION



项目实战 5：互联网分布式爬虫数据采集



猎聘爬虫

开启采集

关键字

web

地区

珠海

优先级

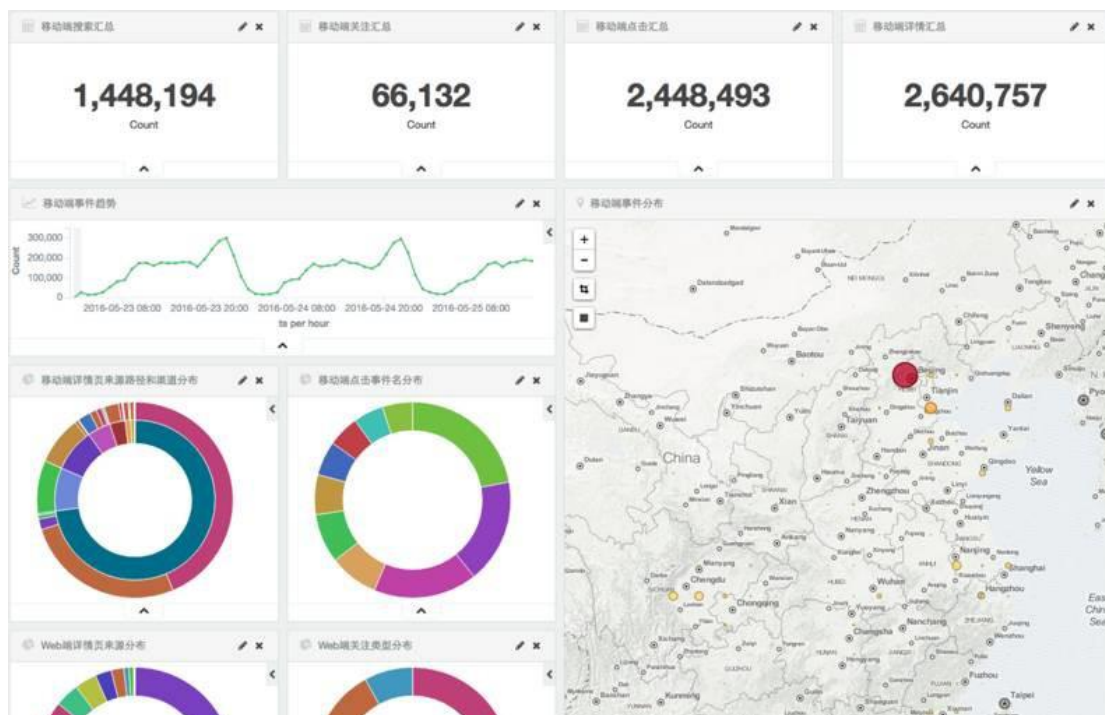
正常

数据列表

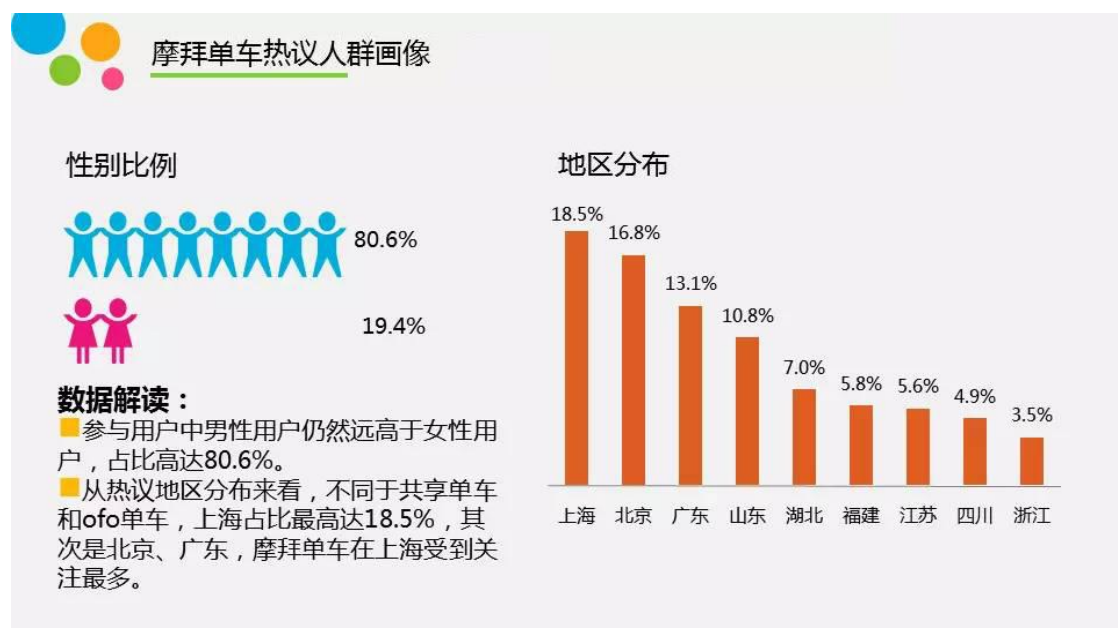
导出数据

公司名称	职位名称	职位薪资	工作地点	年龄要求	经验	职位描述
广东飞企互联科技股份	Web工程师	10-11万	珠海	年龄不限	经验不限	职位描述: 职位描述: .
珠海幸福之家房地产投	IT专员	6-12万	珠海-香洲区	20-40岁	1年以上经验	职位描述: 【岗位职责:
金山软件	Android开发工程师	面议	珠海	年龄不限	2年以上经验	职位描述: 岗位职责: .
格力电器	应用工程师	面议	珠海	年龄不限	2年以上经验	职位描述: 岗位职责:1、
格力电器	系统工程师	面议	珠海	年龄不限	2年以上经验	职位描述: 岗位职责:1、
珠海市长陆工业自动控	软件工程师(武汉)	21-35万	珠海-金湾区	25-55岁	2年以上经验	职位描述: 一、软件项

项目实战 6：互联网舆情大数据挖掘与分析



项目实战 7：共享单车骑行大数据分析



项目实战 8：O2O 外卖点单平台大数据分析



(课程实战项目根据行业技术发展情况及时调整)

十四、结束语

数据就是生产力，大数据金矿的价值已经被越来越多的企业认识到，大数据技术与各个行业的深度融合，催生了大数据分析师这一炙手可热的职业。高薪、轻编程、发展空间广阔、职业周期长久——一个宝贵的职业机遇已经出现。

机者如神，难遇而易失，把握时代机遇、实现职场跃升，成功永远属于眼光独到且不懈奋斗的人。

把握时代机遇，定位职场坐标

职坐标与奋斗者同行