



《大数据分析师行业调查》

|  |  |
| --- | --- |
| **专 业** | **数据科学与大数据** |
| **姓 名** | **卢振浩** |
| **日 期** | **2020/9/1** |

**摘 要**

随着移动互联网、物联网、5G通信网等新兴技术的迅猛发展，数以亿计的网络接入点、联网设备以及网络应用产生的海量数据，给网络故障排查、网络安全保障等带来了极大的挑战，同时也为人们深度挖掘和充分利用网络大数据的巨大价值带来了机遇。大数据分析可以处理海量数据，并从中抽取有价值的潜在知识，帮助决策者发现隐藏的关系和模式，近年来引起了学术界和工业界的广泛关注。

进入大数据时代，数据对于许多行业来说都很重要，有些数据甚至成为企业的资产，掌握了数据，也就掌握了流量。如今的企业需要通过数据分析来实现精细化运营，降低成本、提高效率，“大数据分析师”的职业应运而生。

本文通过行业概述与应届生就业现状、岗位职能介绍、行业需要的主要能力与技术、专业知识在行业中的实践路径距离、就业准备建议来调查介绍大数据分析师这一行业。

**目 录**

[第1章 行业概述与应届生就业现状 - 1 -](#_Toc14478)

[1.1 行业概述 - 1 -](#_Toc8865)

[1.1.1 背景 - 1 -](#_Toc23216)

[1.1.2 行业介绍 - 1 -](#_Toc23216)

[1.1.2.1 大数据分析 - 1 -](#_Toc22557)

[1.1.2.2 大数据分析师 - 1 -](#_Toc22557)

[1.2 应届生就业现状 - 1 -](#_Toc8865)

[1.2.1 就业领域及企业规模 - 2 -](#_Toc23216)

[1.2.2 薪资水平 - 3 -](#_Toc23216)

[第2章 岗位职能介绍 - 4 -](#_Toc14478)

[2.1 分析企业现状 - 4 -](#_Toc8865)

[2.2 临时性分析指标变化 - 4 -](#_Toc8865)

[2.3 专题分析 - 5 -](#_Toc8865)

[2.4 预测及分析 - 5 -](#_Toc8865)

[第3章 行业需要的主要能力与技术 - 5 -](#_Toc14478)

[3.1 SQL数据库 - 5 -](#_Toc8865)

[3.2 统计学基础 - 5 -](#_Toc8865)

[3.3 python - 5 -](#_Toc8865)

[第4章 专业知识在行业中的实践路径举例 - 5 -](#_Toc14478)

[4.1 电视媒体 - 5 -](#_Toc8865)

[4.2 医疗行业 - 6 -](#_Toc8865)

[4.3 能源行业 - 6 -](#_Toc8865)

[4.4 汽车制造 - 6 -](#_Toc8865)

[第5章 就业准备建议 - 7 -](#_Toc14478)

[5.1 打牢基础知识 - 7 -](#_Toc8865)

[5.2 开拓视野广度 - 7 -](#_Toc8865)

[5.3 进行代码积累 - 7 -](#_Toc8865)

[结 论 - 8 -](#_Toc5560)

[参考文献 - 8 -](#_Toc1841)

**第1章 行业概述与应届生就业现状**

## 1.1 行业概述

### 1.1.1 背景

随着移动互联网、物联网、5G通信网等新兴技术的迅猛发展，数以亿计的网络接入点、联网设备以及网络应用产生的海量数据，给网络故障排查、网络安全保障等带来了极大的挑战，同时也为人们深度挖掘和充分利用网络大数据的巨大价值带来了机遇。[1]大数据分析可以处理海量数据，并从中抽取有价值的潜在知识，帮助决策者发现隐藏的关系和模式，近年来引起了学术界和工业界的广泛关注。

网络以快速、大规模和多样化的方式产生数据。Facebook拥有16.5亿用户，每月有10亿活跃用户；微信拥有10多亿用户，其中有７亿活跃用户。[2]随着网络中海量数据的不断积累、大数据分析的迅速发展，隐藏在数据背后的巨大价值也逐渐显现出来。网络运营商充分利用大数据这种宝贵的资源，可以优化网络性能，最大限度地提高网络收益。面对大型数据时，传统数据分析技术存在以下不足。

1. 传统数据分析技术主要处理结构化数据，但大量基于应用程序的数据通常是非结构化的。
2. 数据分析的实现通常局限在一个部门或业务单位，最终的分析结论是基于非常有限的局部角度给出的，而不是全局角度。
3. 传统分析技术主要针对交易数据，对运营数据关注较少，无法实时决策。

大数据分析能够提取比传统数据分析更具洞察力的信息，可以整合各种零星信息，并利用数据挖掘技术深入研究其中各要素之间的关系，在规划与优化、流量预测与控制、安全保障等方面发挥重要作用。例如，与用户相关的完整数据通常分散在不同的业务部门中，大数据分析能够收集分散的数据，从多个角度了解用户行为和偏好，从而描绘出完整的用户画像。

大数据分析的另一重要特征是实时处理，通过大数据分析，运营商可以实时监控基础架构，并做出自主和动态的决策。随着互联网的发展，各应用（如社交网络、物联网、智能电 网等）对现有网络提出了更多要求，如更灵活、更快捷、更安全、更智能。为满足这些需求，可以结合网络中收集的大量数据和分布式高性能计算平台构建基于大数据分析的网络平台，将网络从无视数据管理转换为富有洞察力的上下文感知网络。

1.1.2 行业介绍

1.1.2.1 大数据分析

大数据分析是指对规模巨大的数据进行分析。大数据可以概括为5个V， 数据量大(Volume)、速度快(Velocity)、类型多(Variety)、价值（Value）、真实性(Veracity)。大数据分析涉及到的环节：数据获取、数据存取、数据预处理、数据建模与分析、数据可视化。

1.1.2.1 大数据分析师

大数据分析师是指在不同行业中，专门从事相关数据的收集、整理、分析，并依据数据通过科学算法模型进行行业研究、评估和预测等工作的专项人才。应用行业涉及互联网信息技术企业、科研院校、金融行业、制造业、物流、生物医疗、农业等大数据相关行业。1.2 应届生就业现状

从20世纪90年代起，欧美国家开始大量培养数据分析师，直到现在，对数据分析师的需求仍然长盛不衰，而且还有扩展之势。

据数联寻英发布《大数据人才报告》显示，未来3-5年内大数据人才的缺口将高达100万。根据中国商业联合会数据分析专业委员会统计，未来中国基础性数据分析人才缺口将达到1400万，而在BAT等大型互联网公司的招聘职位里，80%以上都在招大数据人才。进入大数据行业，也成了越来越多人实现职场高薪梦的路径之一。

在美国，大数据分析师每年薪酬高达17.5万美元。在国内，大数据分析师平均薪酬为19970/月

1.2.1 就业领域及企业规模

如今，几乎每个行业都将数据分析作为行业竞争力，以领先于竞争对手，更好地服务客户。从行业招聘数量来看，数据分析师职位缺口主要集中在三大巨头行业：互联网/电子商务、金融及计算机软件；同时非典型数据产业也在迅速崛起。看来数据分析师还是很抢手的嘛，几乎各行各业都离不开。

互联网/电子商务领域交叉销售，再营销，打包服务或个性化产品，客户数据正在以前所未有的方式受到关注，购买模式、搜索历史、行为分析...在越来越多的行业应用下,数据分析师的需求也越来越多。

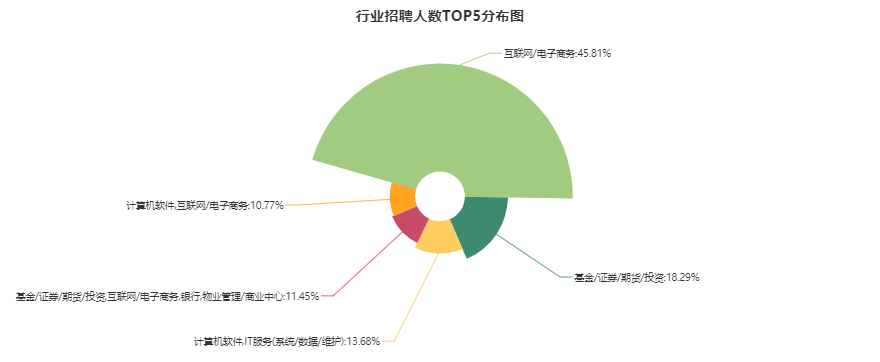


图1.1 行业招聘人数TOP5分布图

从招聘企业规模分布来看，企业规模在100人以上的公司占了70%左右。根据所得数据，其中企业规模在100-499人的公司所占比例最高。数据分析工作离不开企业信息化建设，更需要投入一定的成本。硬性成本与软性成本的投入对于企业规模和资金实力都做出了一定的要求。

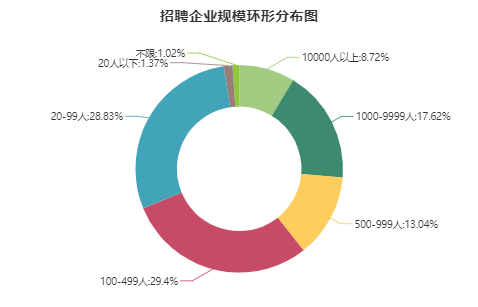


图1.2 招聘企业规模环形分布图

1.2.2 薪资水平

从薪资分布来看，“数据分析师”岗位的平均薪资为 9K以上。并且随着岗位学历的增高，其薪资更是直线上升；拥有“博士”学历的薪资更是达到了18K。

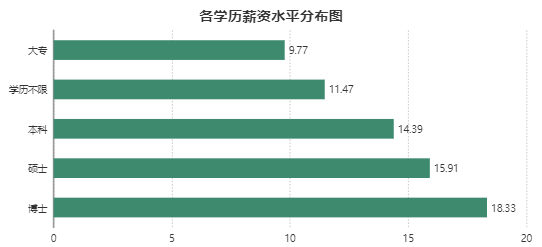


图1.3 各学历薪资水平分布图

从不同企业规模的需求数量与职位来看，企业规模在100-499人的需求数量最多。同时，随着企业规模的扩大，其岗位薪资也在上升。其中企业规模在1000-999人的公司更是给出了最高薪。

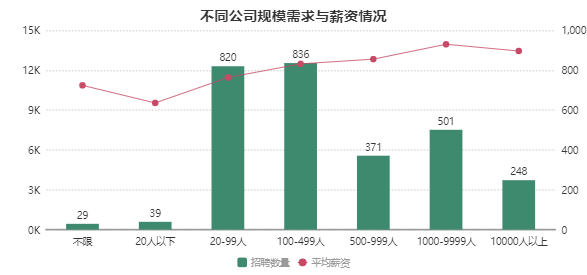


图1.4 不同公司规模需求与薪资情况

1. **岗位职能介绍**

大数据分析师常见工作内容为：1、通过搭建数据指标体系分析企业现状。2、通过检测数据波动变化等随机变量，得到临时性分析指标变化内原的原因。3、通过产生专项事件或周期行的进行专题分析。4、通过符合模型关系和行业属性的大数据参数进行预测及分析。

2.1 分析企业现状

进入大数据时代，数据对于许多行业来说都很重要，有些数据甚至成为企业的资产，掌握了数据，也就掌握了流量。如今的企业需要通过数据分析来实现精细化运营，降低成本、提高效率。

对于企业来说，我们无法用单一的指标来衡量整个企业的运行状况，比如我们只选用生产成本这个指标，就无法得知销售端的状况，只使用销售金额又无法得知今年的利润如何。

为了解决这类问题，系统性思考的能力十分重要，也就是能否站在全局角度，通过严谨的逻辑与业务结合来把握全局。

2.2 临时性分析指标变化

数据分析师的第二块核心工作就是临时数据。无论是现实中还是网络上，不同对象的指数一直都是实时变化的，而数据分析师势必要面对这些不断变化的数据。这些数据往往规模庞大，且产生速度快，难以储存。这些数据中所蕴含的重要信息就会交给大数据分析师进行数据的处理。

除了我们已知的尚未出现的数据，还有一些临时指标往往是因为企业业务本身的变化导致的，这种临时性数据分析能力也是大数据分析师的必备能力。

2.3 专题分析

对公司内的长期业务，大数据分析师也要进行周期性的评估。这种工作往往是固定的，在长期看来不会发生变化的。是大数据分析师的本职工作。

2.4 预测及分析

通过数据进行预测是大数据分析师必备的能力，也是企业中不可或缺的的职位。而模型的准确性与稳定性就十分重要，优化模型以及算法也是大数据分析师需要注意的。

1. **行业需要的主要能力与技术**

3.1 SQL数据库

结构化查询语言简称SQL，是一种数据库查询和程序设计语言，用于存取数据以及查询、更新和管理关系数据库系统；同时也是数据库脚本文件的扩展名。SQL能够帮我们在不同的表中建立联系，也能帮我们截取你需要的特定数据。

3.2 统计学基础

学习统计学基础会让你对数据拥有更好地感知，对数据的变化也更敏感。在一些大数据基础理论中，统计学的基础会帮助我们理解大数据的基础理论。因此统计学是大数据计算科学必不可少的基础。

3.3 python

目前很多程序设计都离不开python，大数据分析也是如此。Python中含有大量的开源程序库，可以帮助我们设计计算程序。

1. **专业知识在行业中的实践路径举例**

4.1 电视媒体

对于体育爱好者，追踪电视播放的最新运动赛事几乎是一件不可能的事情，因为有超过上百个赛事在8000多个电视频道播出。

　　而现在市面上开发了一个可追踪所有运动赛事的应用程序RUWT，它已经可以在iOS和Android设备，以及在Web浏览器上使用，它不断地分析运动数据流来让球迷知道他们应该转换成哪个台看到想看的节目，在电视的哪个频道上找到，并让他们在比赛中进行投票。对于谷歌电视和TiVo用户来说，实际上 RUWT就是让他们改变频道调到一个比赛中。

该程序能基于赛事的紧张激烈程度对比赛进行评分排名，用户可通过该应用程序找到值得收看的频道和赛事。

4.2 医疗行业

在加拿大多伦多的一家医院，针对早产婴儿，每秒钟有超过3000次的数据读取。通过这些数据分析，医院能够提前知道哪些早产儿出现问题并且有针对性地采取措施，避免早产婴儿夭折。

它让更多的创业者更方便地开发产品，比如通过社交网络来收集数据的健康类App。也许未来数年后，它们搜集的数据能让医生给你的诊断变得更为精确，比方说不是通用的成人每日三次一次一片，而是检测到你的血液中药剂已经代谢完成会自动提醒你再次服药。

4.3 能源行业

智能电网现在欧洲已经做到了终端，也就是所谓的智能电表。在德国，为了鼓励利用太阳能，会在家庭安装太阳能，除了卖电给你，当你的太阳能有多余电的时候还可以买回来。通过电网收集每隔五分钟或十分钟收集一次数据，收集来的这些数据可以用来预测客户的用电习惯等，从而推断出在未来2~3个月时间里，整个电网大概需要多少电。有了这个预测后，就可以向发电或者供电企业购买一定数量的电。因为电有点像期货一样，如果提前买就会比较便宜，买现货就比较贵。通过这个预测后，可以降低采购成本。

4.4 汽车制造

当问起汽车的制造过程，大多数人脑子里随即浮现的是各种生产装配流水线和制造机器。然而在福特，在产品的研发设计阶段，大数据就已经对汽车的部件和功能产生了重要影响。

比如，福特产品开发团队曾经对SUV是否应该采取掀背式(即手动打开车后行李箱车门)或电动式进行分析。如果选择后者，门会自动打开、便捷智能，但这种方式会影响到车门开启有限的困恼。此前采用定期调查的方式并没有发现这个问题，但后来根据对社交媒体的关注和分析，发现很多人都在谈论这些问题。

1. **就业准备建议**

5.1 打牢基础知识

大数据分析师最离不开的就是大数据计算的知识，这也是大数据分析师与普通的数据分析师之间的不同之处。因此，无论是计算机的基础知识，还是大数据的专业知识，对这一行业都非常重要。

5.2 开拓视野广度

大数据分析师在工作中遇到的问题常常是多样化的，而不仅仅是计算机行业内部的问题。因此这就需要我们掌握其他领域的专业知识，了解的越多，对于解决问题的帮助越大。多与其他专业的人沟通将是必要的选择。

5.3 进行代码积累

对于任何一个与计算机相关的行业，代码的积累必不可少。而大数据行业更是如此，积累代码就是在积累用来解决问题的模型，模型越多，思路越清晰，解决问题的方式越灵活。进行代码的积累必不可少。

**结 论**

随着大数据、物联网、5G 等技术应用的不断发展，社会对大数据分析师从业人员的需求日益增长，但目前人才数量较少，人才队伍建设亟须加强。而大数据分析师这个行业，时十分适合大数据专业的人才进行的工作。大数据分析师这样的工作，正是大数据学科应用的集大成者。如果在未来想要进入企业，这个工作也将会是优选。

**参考文献**

[1] 冯贵兰, 李正楠, 周文刚. 大数据分析技术在网络领域中的研究综述[J]. 计算机科学, 2019,46(06):1-20.

[2] 大数据分析师：发掘数据力量的多面手\_东南网记者\_\_林先昌[Z].