



书名：大数据会计基础（双色）

ISBN：978-7-5647-9463-7

作者：丁修平

出版社：电子科技大学出版社

定价：52.80元

前言

为了贯彻落实教育部深化“工学结合”人才培养模式，推进课程建设与改革，以培养学生动手操作能力为切入点，通过对会计工作流程分析、提炼、梳理，实现零基础会计基础理论学习对象快速入门。本书根据最新的税收政策标准，对编写内容进行全面更新，以满足财会类专业学生及其他需要学习了解会计基本知识、基本理论和基本技能人士，尤其是企业在岗会计基础培训教学需要。

本书的特点具体表现在以下几个方面。

1.对大数据会计基础进行了介绍，引导学生追寻大数据在会计中应用的基本方向。

2.以单位经济业务内容为依归，以会计工作流程为引导，以任务为抓手展开内容。

本教材突破传统教材对理论的过度阐述，从初学者能够尽快入门并理解所学的内容目的出发，按照企业实际业务流程和必要理论进行项目构建，通过会计认知、会计凭证填制、会计账簿登记、会计报表编制由浅入深，将会计方法、会计要素、会计原则融入操作流程。清晰展现会计日常工作步骤和业务处理的全过程，简单明了，使学习者迅速进阶。

3.明确教学目标与教学对象自主性学习需求，注重内容启发和项目任务有序衔接。

本教材在每一个项目开始，结合教学项目提出学习目标，明确学习要点及项目之间内容的联系，引导学生有目标地学习、领会所学内容项目之间的承接、延续。

4.精选案例，将需要学习掌握的难点融入操作过程，便于内化。内容体系新颖，编排符合学习认知规律。

5.理论必要，通俗易懂，注重理论与实践相结合的指向性要求。

本教材按照新会计准则和财税政策要求编写。编写内容均为最新的会计处理政策、方法，涵盖了会计日常工作中所需知识和技能，帮助学习者快速理解经济业务内容和会计处理基本要求和方法，通过学习教材，能轻松掌握所学并具备动手做账能力。

本书参考借鉴了现行出版教材的优秀成果，在此表示感谢。由于编者水平有限，错漏之处在所难免，恳请读者提出宝贵意见，以便于我们在下次修订中加以完善。

编者

第一篇 大数据基础

项目一

大数据会计基础	3
◎ 任务一 认识大数据	4
◎ 任务二 大数据分析下的会计基础数据特征提取	10
◎ 任务三 大数据分析	12
◎ 任务四 应用 MySQL 数据库建立财务模型	17
◎ 课后练习	20

第二篇 会计基础

项目二

会计与会计工作规范	23
◎ 任务一 认识会计	23
◎ 任务二 会计的基本假定	28
◎ 任务三 会计确认与会计核算方法	30
◎ 任务四 会计的核算信息质量要求	34
◎ 任务五 会计要素	36
◎ 任务六 复式记账	49
◎ 任务七 会计基础工作规范	61
◎ 课后练习	70

项目三

借贷记账法的应用	73
◎ 任务一 资金筹集业务核算	74
◎ 任务二 供应过程业务核算	77
◎ 任务三 生产过程业务核算	81
◎ 任务四 销售过程业务核算	91
◎ 任务五 财务成果形成和分配核算	95
◎ 任务六 资金退出业务核算	101
◎ 课后练习	102

项目四

会计凭证	107
◎ 任务一 会计凭证的意义和种类	107
◎ 任务二 原始凭证的填制和审核	114
◎ 任务三 记账凭证的填制和审核	116
◎ 任务四 会计凭证的传递和保管	118
◎ 课后练习	120

项目五

会计账簿	123
◎ 任务一 会计账簿的意义和种类	123
◎ 任务二 会计账簿的设置和登记	125
◎ 任务三 总分类账户与明细分类账户的平行登记	132
◎ 任务四 对账和结账	138
◎ 任务五 错账的查找和更正	141
◎ 任务六 会计账簿的更换和保管	143
◎ 课后练习	144

项目六

财产清查与账项调整	148
◎ 任务一 财产清查的意义和种类	148
◎ 任务二 财产物资的盘存制度	150
◎ 任务三 财产清查的方法	151
◎ 任务四 财产清查结果的会计处理	154
◎ 任务五 账项调整	155
◎ 课后练习	158

项目七

财务报表	161
◎ 任务一 财务报表概述	161
◎ 任务二 资产负债表编制	165
◎ 任务三 利润表编制	178
◎ 任务四 现金流量表	186
◎ 课后练习	190

项目八

账务处理程序	193
◎ 任务一 账务处理程序的认知	193
◎ 任务二 记账凭证账务处理程序	195
◎ 任务三 汇总记账凭证账务处理程序	196
◎ 任务四 科目汇总表账务处理程序	200
◎ 课后练习	213

习题答案

.....	216
-------	-----

参考文献

.....	222
-------	-----

第一篇 大数据基础

项目一 大数据会计基础

项目一

大数据会计基础

知识目标

- (1) 认识大数据。
- (2) 大数据分析下的会计基础数据特征提取。
- (3) 大数据分析。

技能目标

- (1) 应用 MySQL 数据库建立财务模型。



互联网大大改变了个人的生活方式。在国外出现了雅虎、谷歌、脸谱、推特、亚马逊等国际知名互联网企业，在国内出现了新浪、搜狐、百度、腾讯、阿里巴巴等知名互联网企业。互联网为人们提供了新闻、搜索、即时消息、电子商务等不计其数的免费（或称“廉价”）的服务，人们在享受互联网提供的便捷、廉价服务的同时，互联网服务平台也悄悄地记录下个人的特征和行为信息。

以社交网络系统为例，通过大数据分析，可以发现个人的社会关系网络及其强弱程度。对于个人，可以发现多年不联系的同学、老乡、战友、同事或者具有共同兴趣爱好的朋友等；对于企业，可以通过社交关系分析发现所需专业人才；对于公共安全管理部，可以通过社交关系分析预测重点关注人员的犯罪倾向，并提前采取措施。

电子商务消除了商品交易过程中的信息不对称现象，有效地匹配了市场需求和供给，提高了商品购买的便捷性，因此又被称作“无摩擦经济”“眼球经济”“新经济”。电子商务平台记录了商品交易信息，包括商家的采购、库存、销售、服务等记录以及买家浏览、搜索、选择、购买、咨询、投诉、建议等记录，形成了电子商务大数据。

大数据与会计基础，你希望知道：

- (1) 大数据是什么？它有什么特征？
- (2) 每天产生的数据规模是多少？为什么如此重要？
- (3) 大数据会不会改变会计核算基础工作内容？
- (4) 大数据对会计信息质量有哪些改变？
- (5) 如何学好大数据会计？



任务一 认识大数据

一、会计的产生与发展

(一) 大数据的概念与特点

大数据是指需要新处理模式才能具有更强的决策力、洞察力和流程化能力的海量、高增长率和多样化信息资产，具有大量、高速、多样、价值密度、真实性等特点。

(二) 大数据的产生与发展

大数据概念最早的提出者现已无法考证。1980年，未来学家托夫勒在其所著的《第三次浪潮》中就提到“大数据”一词。

互联网领域的公司最早重视数据资产的价值，最早从大数据中淘金，并且引领大数据的发展趋势。随着互联网技术发展，信息基础设施持续完善，包括网络带宽的持续增加、存储设备功能不断提升，为大数据的存储和传播提供了良好的物质基础。云计算为大数据的集中管理和分布式访问提供了必要的场所和分享的渠道。

1. 国外产生与发展情况

2005年Hadoop项目诞生。Hadoop最初只是雅虎公司用来解决网页搜索问题的一个项目，后来因其技术的高效性，被Apache Software Foundation公司引入并成为开源应用。

Hadoop本身不是一个产品，而是由多个软件产品组成的一个生态系统，这些软件产品共同实现全面功能和灵活的大数据分析。从技术上看，Hadoop由两项关键服务构成：采用Hadoop分布式文件系统（HDFS）的可靠数据存储服务，以及利用一种叫作MapReduce技术的高性能并行数据处理服务。这两项服务的共同目标是，提供一个使结构化和复杂数据的快速、可靠分析变为现实的基础。

2008年年末，“大数据”得到部分美国知名计算机科学研究人员的认可，业界组织计算社区联盟（Computing Community Consortium），发表了一份有影响力的白皮书《大数据计算：在商务、科学和社会领域创建革命性突破》。它使人们的思维不仅局限于数据处理的机器，并提出，大数据真正重要的是新用途和新见解，而非数据本身。此组织可以说是最早提出大数据概念的机构。

2009年印度政府建立了用于身份识别管理的生物识别数据库，联合国全球脉冲项目已研究了对如何利用手机和社交网站的数据源来分析预测从螺旋价格到疾病暴发之类的问题。

2009年，美国政府通过启动<http://Data.gov>网站的方式进一步开放了数据的大门，这个网站向公众提供各种各样的政府数据。该网站超过4.45万的数据集被用于保证一些网站和智能手机应用程序来跟踪从航班到产品召回再到特定区域内失业率的信息，这一行动激发了从肯尼亚到英国范围内的政府相继推出类似举措。

2009年，欧洲一些领先的研究型图书馆和科技信息研究机构建立了伙伴关系，致力于改善在互联网上获取科学数据的简易性。



2010年2月,肯尼斯·库克尔在《经济学人》上发表了长达14页的大数据专题报告《数据,无所不在的数据》。库克尔在报告中提到:世界上有着无法想象的巨量数字信息,并以极快的速度增长。从经济界到科学界,从政府部门到艺术领域,很多方面都已经感受到了这种巨量信息的影响。科学家和计算机工程师已经为这个现象创造了一个新词汇——大数据。库克尔也因此成为最早洞见大数据时代趋势的数据科学家之一。

2011年2月,IBM的沃森超级计算机每秒可扫描并分析4TB(约2亿页文字量)的数据量,并在美国著名智力竞赛电视节目《危险边缘》上击败两名人类选手而夺冠。后来纽约时报认为这一时刻为一个“大数据计算的胜利。”

2011年5月,全球知名咨询公司麦肯锡的全球研究院(MGI)发布了一份报告——《大数据:创新、竞争和生产力的下一个新领域》,大数据开始备受关注,这也是专业机构第一次全面地介绍和展望大数据。报告指出,大数据已经渗透到当今每一个行业和业务职能领域,成为重要的生产因素。人们对于海量数据的挖掘和运用,预示着新一波生产率增长和消费者盈余浪潮的到来。报告还提到,“大数据”源于数据生产和搜集的能力和速度的大幅提升——由于越来越多的人、设备和传感器通过数字网络连接起来,产生、传送、分享和访问数据的能力也得到彻底变革。

2012年,大数据领域发展成果丰硕。

2012年,《大数据时代》出版,作者V.迈尔-舍恩伯格被誉为“大数据商业应用第一人”。

2012年1月,瑞士达沃斯召开的世界经济论坛上,大数据是主题之一,会上发布的报告《大数据,大影响》(*Big Data, Big Impact*)宣称,数据已经成为一种新的经济资产类别,就像货币或黄金一样。

2012年3月,美国奥巴马政府在白宫网站发布了《大数据研究和发展倡议》,这一倡议标志着大数据已经成为重要的时代特征。2012年3月22日,奥巴马政府宣布将2亿美元投资到大数据领域,是大数据技术从商业行为上升到国家科技战略的分水岭,在次日的电话会议中,政府对数据的定义为“未来的新石油”,大数据技术领域的竞争,事关国家安全和未来。并表示,国家层面的竞争力将部分体现为一国拥有数据的规模、活性以及解释、运用的能力;国家数字主权体现对数据的占有和控制。数字主权将是继边防、海防、空防之后,另一个大国博弈的空间。

2012年4月,美国软件公司Splunk于19日在纳斯达克成功上市,成为第一家上市的大数据处理公司。鉴于美国经济持续低迷、股市持续震荡的大背景,Splunk首日的突出交易表现尤其令人们印象深刻,首日即暴涨了一倍多。Splunk是一家领先的提供大数据监测和分析服务的软件提供商,成立于2003年。Splunk成功上市促进了资本市场对大数据的关注,同时也促使IT厂商加快大数据布局。

2012年7月,联合国在纽约发布了一份关于大数据政务的白皮书,总结了各国政府如何利用大数据更好地服务和保护人民。这份白皮书举例说明在一个数据生态系统中,个人、公共部门和私人部门各自的角色、动机和需求,例如:通过对价格关注和更好服务的渴望,个人提供数据和众包信息,并对隐私和退出权力提出需求;公共部门出于改善服务、提升效益的目的,提供了诸如统计数据、设备信息、健康指标及税务和消费信息等,并对隐私和退出权力提出需求;私人部门出于提升客户认知和预测趋势目的,提



供汇总数据、消费和使用信息，并对敏感数据所有权和商业模式更加关注。白皮书还指出，人们如今可以使用极大、丰富的数据资源，包括旧数据和新数据，来对社会人口进行前所未有的实时分析。联合国还以爱尔兰和美国的社交网络活跃度增长可以作为失业率上升的早期征兆为例，表明政府如果能合理分析所掌握的数据资源，将能“与数俱进”，快速应变。

2012年7月，为挖掘大数据的价值，阿里巴巴集团在管理层设立“首席数据官”一职，负责全面推进“数据分享平台”战略，并推出大型的数据分享平台——聚石塔，为天猫、淘宝平台上的电商及电商服务商等提供数据云服务。随后，阿里巴巴董事局主席马云在2012年网商大会上发表演讲，称从2013年1月1日起将转型重塑平台、金融和数据三大业务。马云强调：“假如我们有一个数据预报台，就像为企业装上了一个GPS和雷达，你们出海将会更有把握。”因此，阿里巴巴集团希望通过分享和挖掘海量数据，为国家和中小企业提供价值。此举是国内企业最早把大数据提升到企业管理层高度的一次重大里程碑。阿里巴巴也是最早提出通过数据进行企业数据化运营的企业。

2013年，IBM公司提出大数据的5V特征，即Volume（数据量巨大）、Velocity（高速及时有效分析）、Variety（种类和来源的多样化）、Value（价值密度低，商业价值高）、Veracity（数据的真实有效性）。

2014年4月，世界经济论坛以“大数据的回报与风险”为主题发布了《全球信息技术报告（第13版）》。报告认为，在未来几年中针对各种信息通信技术的政策甚至会显得更加重要。在接下来将对数据保密和网络管制等议题展开积极讨论。全球大数据产业的日趋活跃，技术演进和应用创新的加速发展，使各国政府逐渐认识到大数据在推动经济发展、改善公共服务，增进人民福祉，乃至保障国家安全方面的重大意义。

2014年5月，美国白宫发布了2014年全球大数据白皮书的研究报告《大数据：抓住机遇、守护价值》。报告鼓励使用数据以推动社会进步，特别是在市场与现有的机构并未以其他方式来支持这种进步的领域；同时，也需要相应的框架、结构与研究，来帮助保护美国人对于保护个人隐私、确保公平或是防止歧视的坚定信仰。

在重视发展科技的印度，大数据技术也已成为信息技术行业的“下一个大事件”。目前，不仅印度的小公司纷纷涉足大数据市场淘金，一些外包行业巨头也开始进军大数据市场，试图从中分得一杯羹。2016年，印度全国软件与服务企业协会预计，印度大数据行业规模在3年内将到12亿美元，是当前规模的6倍，同时还是全球大数据行业平均增长速度的两倍。印度毫无疑问是美国亦步亦趋的好学生。在数据开放方面，印度效仿美国政府的做法，制定了一个一站式政府数据门户网站data.gov.in，把政府收集的所有非涉密数据集中起来，包括全国的人口、经济和社会信息。

英国在2017年议会期满前，开放有关交通运输、天气和健康方面的核心公共数据库，并在五年内投资1000万英镑建立世界上首个“开放数据研究所”，政府将与出版行业等共同尽早实现对得到公共资助产生的科研成果的免费访问，英国皇家学会也在考虑如何改进科研数据在研究团体及其他用户间的共享和披露；英国研究理事会将投资200万英镑建立一个公众可通过网络检索的“科研门户”。

截至2018年11月，美国政府公共数据共享网站（Data.gov）上公共数据创新应用案



例所属领域分布情况，其中集中在商业贸易、农业、健康医疗等领域的创新应用最多，分别占比 18.9%、13.3%、10.0%。

目前，欧盟在大数据方面的活动主要涉及四方面内容：研究数据价值链战略因素；资助“大数据”和“开放数据”领域的研究和创新活动；实施开放数据政策；促进公共资助科研实验成果和数据的使用及再利用。

2. 国内产生与发展情况

2011 年 12 月，中华人民共和国工业和信息化部（简称“工信部”）发布的物联网“十二五”规划上，把信息处理技术作为四项关键技术创新工程之一提出来，其中包括了海量数据存储、数据挖掘、图像视频智能分析，这都是大数据的重要组成部分。

2014 年，“大数据”首次出现在当年的《政府工作报告》（以下简称《报告》）中。《报告》中指出，要设立新兴产业创业创新平台，在大数据等方面赶超先进，引领未来产业发展。“大数据”旋即成为国内热议词汇。

2015 年 8 月，国务院正式印发《促进大数据发展行动纲要》（以下简称《纲要》），《纲要》明确，推动大数据发展和应用，在未来 5~10 年打造精准治理、多方协作的社会治理新模式，建立运行平稳、安全高效的经济运行新机制，构建以人为本、惠及全民的民生服务新体系，开启大众创业、万众创新的创新驱动新格局，培育高端智能、新兴繁荣的产业发展新生态。提出建设政府数据资源共享开放、国家大数据资源统筹发展、大数据关键技术及产品研发与产业化、大数据产业支撑能力提升、网络和大数据安全保障、政府治理、公共服务、工业和新兴产业、现代农业、万众创新等 10 个大数据工程，加快建设数据强国。标志着大数据正式上升为国家战略。

2016 年 3 月，《中华人民共和国国民经济和社会发展第十三个五年规划纲要》提出，实施国家大数据战略，把大数据作为基础性战略资源，全面实施促进大数据发展行动，加快推动数据资源共享开放和开发应用，助力产业转型升级和社会治理创新。大数据技术的战略意义在于对庞大的、含有意义的数据进行专业化处理，具有实时交互式的查询效率和分析能力。

国务院总理李克强 2016 年 5 月 25 日在出席中国大数据产业峰会暨中国电子商务创新发展峰会时指出，坚持创新驱动发展战略，制定实施网络强国战略、国家大数据战略、“互联网+”行动计划、《中国制造 2025》等，出台了一系列重大举措，在促进创新和更好发挥人力资源优势方面的效应正在显现。以大数据为代表的创新意识和传统产业长期孕育的工匠精神相结合，使新旧动能融合发展，并带动改造和提升传统产业，有力推动虚拟世界和现实世界融合发展，打造中国经济发展的“双引擎”。

习近平总书记在中共十九大报告中指出，要“推动互联网、大数据、人工智能和实体经济深度融合，在中高端消费、创新引领、绿色低碳、共享经济、现代供应链、人力资本服务等领域培育新增长点、形成新动能”。

2018 年，我国数字政府建设进程明显加快，各级政府正积极挖掘互联网和大数据技术在社会治理中的作用，优化社会态势感知、畅通沟通渠道、辅助决策施政。展望未来，政府门户网站、政务服务 APP、公共服务公众号等将成为民主协商的重要平台，互联网和大数据技术的融合能给社会个体提供有效的政治参与渠道，社会治理的主体将由一元



主体转向多元主体变，治理方式将从单向控制、代议互动转向数字协商。

截至 2018 年 12 月，我国网民规模达 8.29 亿，普及率达 59.6%，较 2017 年底提升 3.8 个百分点，全年新增网民 5653 万。我国手机网民规模达 8.17 亿，网民通过手机接入互联网的比例高达 98.6%。

2021 年新的政策是《十四五规划》和《十四五大数据产业发展规划》，规划突出数据再数字经济中的关键作用，加强数据要素市场规划建设、重视大数据相关基础设施建设。

2021 年出台了两部法律，分别是 2021 年 6 月的《数据安全法》和 2021 年 8 月的《个人信息保护法》，结合 2016 年 11 月的《网络安全法》，形成了大数据法律的三驾马车。

值得一提的是，《数据安全法》明确政务数据以公开为原则，不公开为例外的基本理念。因此可以预见，政府主导的数据资源共享交换平台会是十四五期间建设的重点。

综上所述，大数据发展可以分成三个阶段，每个阶段都有自己的特征和功能。

（1）大数据 1.0 阶段——事务化阶段

数据库管理和数据仓储被视为大数据第一阶段的核心组成部分。其使用数据库查询（SQL）、在线分析处理和标准报告工具等众所周知的技术，为现代数据分析奠定了基础。

（2）大数据 2.0 阶段——网络化阶段

进入 21 世纪以来，互联网开始提供独特的数据搜集和数据分析的机会。随着网络流量和在线经营业务的发展，雅虎、亚马逊和易贝等公司开始通过分析点击率、IP 位置数据和搜索日志来分析客户行为。前期主要是在超文本传输协议和万维网导致半结构化和非结构化数据大量增加，当前，社交媒体数据（视频数据存储）的极大增长，极大加剧了对新型工具技术分析和分析技术的需求，以便能够从这些非结构化数据中提取出有意义的信息。

（3）大数据 3.0 阶段——智能化阶段

虽然基于 Intel 网的非结构化数据内容居多，但是，移动设备、物联网兴起使数据产生量发生巨大飞跃，每天，这些 ZB 级海量数据为运输业、智慧城市设计、现代医疗以及电子商务发展带来巨大商机。而利用人工智能和机器学习技术变得不可或缺，大数据分析进入了一个全新时代，这个时代更加依赖于算法。

（三）大数据对会计的影响

没有大数据分析，公司就会盲目，就像高速公路上的鹿一样（杰弗里·摩尔）。大数据所涉及的数据量规模巨大，它的特点可以概括为多样性、高速性、易变性、大量性、结构复杂性以及真实性。因此，大数据具有为企业提供过去有用信息的作用。它不仅仅体现在数量多，更重要的是数据蕴含的价值很大。大数据对会计的影响不仅体现在数据的形式和规模上，也体现在大数据的处理技术上。

1. 大数据对会计数据来源的影响

传统的会计信息的数据类型应该是结构型的，它易于分析、管理，甚至是采纳；而大数据的类型大多数是非结构型的，就会对会计信息的来源造成一定的影响，也会增加采集成本。



2. 大数据对会计计量方法的影响

目前,会计计量属性是以历史成本为基础的基本框架。企业对于房屋设备等固定资产以及无形资产等都是按照原始成本计价。但是随着经济和信息技术的发展,特别是大数据时代的到来,历史成本逐渐显现出各种弊端。一是历史成本是静态的,它反映的是过去的信息,但是经济社会是在不断变化的,历史成本无法提供最实时的信息;二是提高公允价值计量的可靠性和可操作性。与历史成本计量相比,公允价值有着不可比拟的优越性,公允价值具有的优点是显示出企业的经济实质,能够给利用企业信息的人更多及时有效的决策信息;实现收入与成本、费用的适当比例;保障资本结构,稳定资产负债观。但是,公允价值的运用,受获取公允价值的信息的时效影响,也可能存在“违背公平交易”的一些无法避免的情况。但大数据平台可以为企业提供财务信息共享,实现数据的实时搜集、传播、处理和存储。这样就大大增加了财务数据的可获得性,便于企业获取与确定公允价值的相关变量。

3. 大数据对会计工作模式的影响

大数据条件下,要求企业的财务会计必须重新审视财务会计工作。随着企业的快速发展,企业的财务管理模式也需要转变,以适应大数据环境下企业的发展过程,为企业的发展决策提供准确的信息。首先是信息披露质量的提升。在这个方面企业需要提供准确的财务指标进行考量,同时也要确保数据的可靠性。大数据对提高会计信息处理的效率与质量,准确推进财务工作的进行,并且避免存在一些误差具有重大意义。例如,通过制定新的会计报表附注内容要求,将有关企业的大数据纳入附注要求,可以使信息关联方获得更丰富的企业财务非货币化资讯。

4. 大数据对会计目标定位的影响

会计目标定位不应当仅仅是用于反应报表层面的一些信息,还应该在会计工作中实现有助于促进会计信息多样化的使用。现代企业的财务工作已经不再是过去要求比较低的地步,而是要求突破记账这一种目的,如何不断地进行数据的分析、挖掘数据背后的意义、得出更直观的分析效果才是财务工作应该发展的方向。大数据使得企业、财务数据更加融合,会计人员要如何面对这种趋势,才能提升财务信息质量,为企业经营管理提供更多的参考。

5. 大数据会计基础工作处理方法带来变革

大数据信息技术的应用有着不可替代的优越性。

会计基础工作主要是对涉及会计原始数据的确认、计量、输入、储存、处理、传递、反馈、输出、发布问题,由于大数据不同于以往单纯数据的形态且与现已使用的会计原则存在不一致情况,可能会发生会计基础数据在处理内容、技术、方法上出现新的要求,对会计基础工作带来一系列影响。利用计算机计算速度的优势,能够在数据整合、信息检索等工作上极大地节约人力,减轻财会工作人员的工作负担,利用该技术能够准确地对数据进行查找、修改、综合分析,将更多精力花费在数据分析工作上,从而提升工作的质量。这是传统的工作方式无法做到的,尤其在数据量巨大的今天。

6. 大数据处理的结果使决策更加快速和精准

大数据时代的到来,经济生活领域发生了巨大的变化。基于云计算模式的数据处理,



用户不再需要具体了解内部复杂的算法分析,只需要选择合适的方法来处理数据,就可以最终得到数据处理的结果。借助大数据处理技术,财务会计工作能够提供更详细、直观的市场信息,从而直接利用这些结果进行决策,使得决策更加快速和精准。

7. 大数据对会计工作提出更高要求

当前,从企业的角度来说,应当从硬件、系统以及人员方面做好发展大数据技术的环境。首先,应当建立能够满足大数据技术运行的系统。一个优秀的数据处理系统能够极大地促进财务会计的工作效率,能够让财务会计工作人员更加快速地进行数据的处理。其次,企业应当重视财务会计工作者工作能力的提升,通过定期的培训学习,逐渐提升财务会计工作者的工作能力,提升财务人员应用大数据技术解决问题的能力,充分发挥大数据技术的优势。

任务二 大数据分析下的会计基础数据特征提取

一、大数据技术

会计数据和信息的加工处理流程,即确认、计量、输入、储存、处理、传递、反馈、输出、发布。

大数据技术主要有以下几种。

1. Hadoop

Hadoop 诞生于 2005 年,其最初只是雅虎公司用来解决网页搜索问题的一个项目,后来因其技术的高效性,被 Apache Software Foundation 公司引入并成为开源应用。Hadoop 本身不是一个产品,而是由多个软件产品组成的一个生态系统,这些软件产品共同实现全面功能和灵活的大数据分析。从技术上看,Hadoop 由两项关键服务构成,采用 Hadoop 分布式文件系统(HDFS)的可靠数据存储服务,以及利用一种叫作 MapReduce 技术的高性能并行数据处理服务。

2. Hive

Hive 是一种建立在 Hadoop 文件系统上的数据仓库架构,并能对存储在 HDFS 中的数据进行分析和处理。它最初是应 Facebook 每天产生的海量新兴社会网络数据进行管理和机器学习的需求而产生和发展的。后来其他公司也开始使用和开发 Apache Hive,例如 Netflix、亚马逊等。

3. Storm

Storm 是 Twitter 开源的分布式实时大数据处理框架,被业界称为实时版 Hadoop,主要由 Clojure 编程语言编写。Storm 也为实时计算提供了一些简单高效的原子,而且 Storm 的 Trident 是基于 Storm 原子更高级的抽象框架,类似于基于 Hadoop 的 Pig 框架,让开发更加便利和高效。

随着越来越多的场景对 Hadoop 的 MapReduce 高延迟无法容忍,比如网站统计、推荐系统、预警系统、金融系统(高频交易、股票)等,大数据实时处理解决方案(流计算)的应用日趋广泛,目前已是分布式技术领域最新爆发点,而 Storm 更是流计算技术



中的佼佼者和主流。

Storm 应用场景主要为推荐系统（实时推荐，根据下单或加入购物车推荐相关商品）、金融系统、预警系统、网站统计（实时销量、流量统计，如淘宝“双 11”效果图）、交通路况实时系统等。具有适用场景广泛、可伸缩性高、保证无数据丢失等特点。

4. Python

Python 的创始人为荷兰人吉多·范罗苏姆。由于 Python 语言的简洁性、易读性以及可扩展性，众多开源的科学计算软件包都提供了 Python 的调用接口，例如著名的计算机视觉库 OpenCV、三维可视化库 VTK、医学图像处理库 ITK。而 Python 专用的科学计算扩展库就更多了，十分经典的科学计算扩展库有 NumPy、SciPy 和 Matplotlib，它们分别为 Python 提供了快速数组处理、数值运算以及绘图功能。因此 Python 语言及其众多的扩展库所构成的开发环境十分适合工程技术、科研人员处理实验数据、制作图表，甚至开发科学计算应用程序。

二、数据分析思维的步骤

建立数据分析思维，提升数据分析思维能力，需要遵循以下三个步骤。

1. 深刻理解业务

不同的企业，不同的业务岗位，大家对数据的具体需求虽然不一样，但最根本的是要通过数据了解企业的状况，解决所遇到的问题，为企业经营决策提供依据。

2. 建立数据采集思路

随着大数据技术的不断发展，大数据的技术体系已经趋于成熟，由于大数据技术体系比较庞大，所以采集大数据的时候首先应该根据自身的需求，找到一个适合的切入点。从而避免过多无效数据对存储资源的浪费，提高数据采集工作效率。

大数据采集思路一般可以通过构建思维导图来实现。

3. 提供优化服务

分解业务关键指标，提供优化策略。

例如，税收征管部门对税收分析是以税收经济现象的数量为研究对象的认识活动。深化税收数据分析，可以帮助我们认识税收数量特征、深化理解税收、推断和预测税收收入，从而掌握税收管理的主动权。

税收分析指标是用于分析纳税人税收风险的计算公式及其属性标识，若干指标集合由税收分析预测模型识别。税收风险指标按与税款的关联度，分为税种类、行为类和特定事项类。税种类风险指标主要包括企业增值税类、企业所得税类、土地增值税类、个人所得税类、房产税、土地使用税类、印花税类、契税、城建税及附加类等；行为类风险包括发票行为类、登记类、申报类、税款征收类和管理认定类等；特定事项类风险主要包括非货币性交易、债务重组、拆迁、搬迁、破产、合并、分立、土地房产转让等。通过分析对比单户企业与本地同行业企业的税收贡献率，对低于同行业平均贡献率的企业，提示预警信息并深入分析企业是否存在多计生产成本、多结转销售成本、多计期间费用、扩大税前扣除范围或不计、少计销售收入等问题。将计算出的税率与企业以往年度或本地同行业水平相比，分析判断企业是否存在以上问题。



在税收分析管理的过程中，税收分析识别必须通过构建一套指标和模型来进行。在模型构建过程中，需根据指标对模型管理的重要性设置不同权重，尤其是关键指标的构建。通过设置多层次指标，如一级指标、二级指标、三级指标等，动态管理实际情况。此外，指标体系自身的建设也是至关重要的，如：指标名称、指标功能、取数口径、比对方式、指标在模型中的权重、预警值的计算等。

再如，企业会计部门对销售额的划分可以按照产品品种结构分类，也可以按照收入地域来源分类，还可以按照新老客户的比例分类，甚至还可以按照赊销或现销进行分类，按照业务主次进行分类……从而得到不同的使用目的。

任务三 大数据分析

大数据已经成为当前热门的 IT 行业的词汇，它必将对会计信息使用带来深刻影响。会计大数据挖掘，有赖于数据仓库、数据安全、数据分析、数据挖掘等围绕大数据的商业价值开发行为。随着大数据时代的来临，大数据分析也应运而生。

一、数据分析

数据分析是指用适当的统计分析方法对搜集来的大量数据进行分析，为提取有用信息和形成结论而对数据加以详细研究和概括总结的过程。数据分析有狭义和广义之分，如图 1-1 所示。狭义的数据分析是指根据分析目的，采用对比分析、分组分析、交叉分析和回归分析等分析方法对搜集来的数据进行处理与分析提取有价值的信息，发挥数据的作用，并得到一个统计量结果的过程。广义的数据分析是指针对搜集来的数据运用基础探索、统计分析、深层挖掘等方法，发现数据中有用的信息和未知的规律与模式，进而为下一步的业务决策提供理论与实践依据。广义的数据分析就包含了数据挖掘。

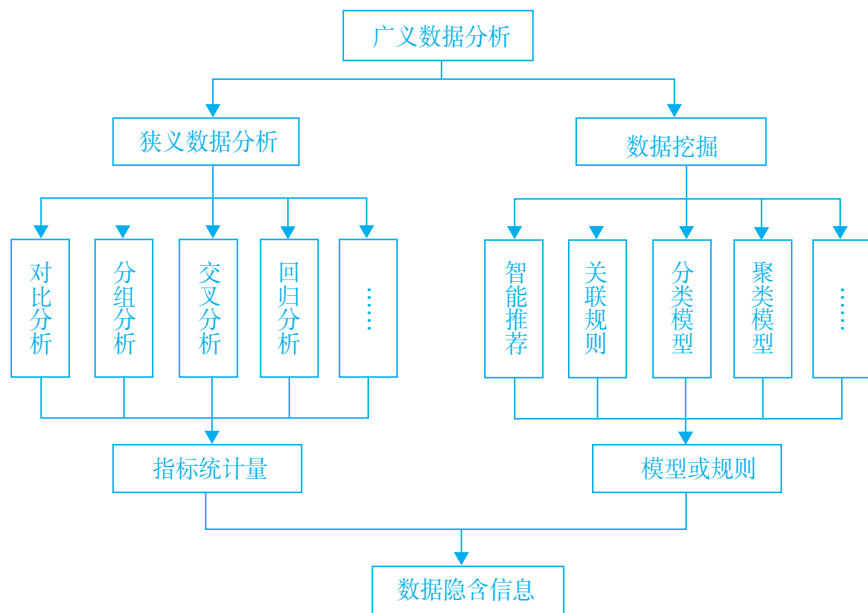


图 1-1 大数据分析



二、大数据分析内容

1. 可视化分析 (Analytic Visualizations)

不管是对数据分析专家还是普通用户，数据可视化是数据分析工具最基本的要求。可视化可以直观地展示数据，让数据自己说话，让观众听到结果。

2. 数据挖掘算法 (Data Mining Algorithms)

可视化是给人看的，数据挖掘就是给机器看的。集群、分割、孤立点分析还有其他的算法让我们深入数据内部，挖掘价值。这些算法不仅要处理大数据的量，也要处理大数据的速度。

3. 预测性分析能力 (Predictive Analytic Capabilities)

数据挖掘可以让分析员更好地理解数据，而预测性分析可以让分析员根据可视化分析和数据挖掘的结果做出一些预测性的判断。

4. 语义引擎 (Semantic Engines)

我们知道由于非结构化数据的多样性带来了数据分析的新的挑战，我们需要一系列的工具去解析、提取、分析数据。语义引擎需要被设计成能够从“文档”中智能提取信息。

5. 数据质量和数据管理 (Data Quality and Master Data Management)

数据质量和数据管理是一些管理方面的最佳实践。通过标准化的流程和工具对数据进行处理可以保证一个预先定义好的高质量的分析结果。

6. 数据存储，数据仓库 (Data Storage, data warehouse)

数据仓库是为了便于多维分析和多角度展示数据按特定模式进行存储所建立起来的关系型数据库。在商业智能系统的设计中，数据仓库的构建是关键，是商业智能系统的基础，承担对业务系统数据整合的任务，为商业智能系统提供数据抽取、转换和加载(ETL)，并按主题对数据进行查询和访问，为联机数据分析和数据挖掘提供数据平台。

三、数据分析的流程

数据分析是基于搜集来的数据，应用数学、统计和计算机等技术抽取出数据中的有用信息，进而为决策提供依据和指导方向。数据分析的流程如下。

1. 需求分析

数据分析中的需求分析也是数据分析环节的第一步和最重要的步骤之一，决定了后续的分析的方向、方法。

2. 数据获取

数据是数据分析工作的基础，是指根据需求分析的结果提取，搜集数据。

3. 数据预处理

数据预处理是指对数据进行数据合并，数据清洗，数据变换和数据标准化，数据变换后使得整体数据变得干净整齐，可以直接用于分析建模这一过程的总称。

4. 分析与建模

分析与建模是指通过对比分析、分组分析、交叉分析、回归分析等分析方法和聚类、分类、关联规则、智能推荐等模型与算法发现数据中的有价值信息，并得出结论的过程。



5. 模型评价与优化

模型评价是指对已经建立的一个或多个模型，根据其模型的类别，使用不同的指标评价其性能优劣的过程。

6. 部署

部署是指通过了正式应用数据分析结果与结论应用至实际生产系统的过程。

数据分析是一个探索性的过程，通常从特定的问题开始，而数据分析和数据可视化这两个术语密不可分。在实际处理数据时，数据分析先于可视化输出，而可视化分析又是呈现有效分析结果的一种好方法。

四、数据可视化

数据可视化是关于数据视觉表现形式（即一种以某种概要形式抽提出来的信息，包括相应信息单位的各种属性和变量）的科学技术研究。它主要是借助图形化手段，清晰有效地传达与沟通信息（即数据的可视化展示），有助于确定需要进一步调查的异常值、差距、趋势和有趣的数据点。有效的可视化可显著减少受众处理信息和获取有价值见解所需的时间，既是一个化繁为简的过程，也是一个表达观点、发现联系的过程，如图 1-2 所示。

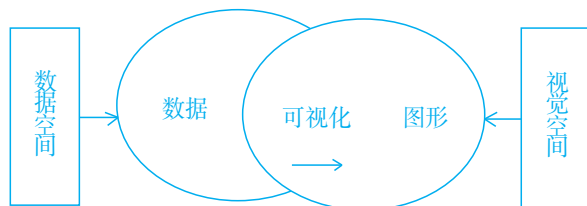


图 1-2 数据可视化

数据聚焦于解决数据的采集、清理、预处理、分析和挖掘。

图形聚焦于解决对光学图像进行接收、提取信息、加工变换、模式识别及存储显示。

可视化聚焦于解决将数据转换成图形，并进行交互处理用于展现分析的前端开源工具有 JasperSoft、Pentaho、Spagobi、Openi、Birt 等。

五、数据分析与可视化常用工具

1. Microsoft Excel

Excel 是大家熟悉的电子表格软件，已被广泛使用了很多年，如今甚至有很多的数据只能以 Excel 表格的形式获取，但是它的局限在于它一次性所能处理的数据量，而且除非通晓 VBA 这个 Excel 内置的编程语言，否则针对不同数据集来绘制一张图表将是一件极其烦琐的事。

2. R 语言

R 语言是集统计分析与图形显示于一体的用于分析、绘图的语言和操作环境的软件，是属于 GNU 系统的一个自由、免费、源代码开放的软件，是一个用于统计计算和统计制图的优秀工具。



3. Python 语言

Python 是一种动态的、面向对象的、解释型脚本语言，起初被用于编写自动化脚本，后随着版本迭代及功能升级，目前也可以用于单独开发，是一门跨平台的脚本语言（Python 规定了一个 Python 语法规则，实现了 Python 语法的解释程序就成为 Python 的解释器）。Python 代码具有简洁性、易读性和易维护性等优点。

4. JavaScript

JavaScript（缩写为 JS）是一种高级的、多范式、解释型的编程语言，是一门基于原型、函数先行的语言，它支持面向对象编程、命令式编程以及函数式编程。它提供语法来操控文本、数组、日期以及正则表达式，不支持 I/O（比如网络、存储和图形等），但可以由它的宿主环境提供支持，是一种基于对象和事件驱动并具有相对安全性的客户端脚本语言。

5. PHP

PHP（“超文本预处理器”）是一种通用开源脚本语言。语法吸收了 C 语言、Java 和 Perl 的特点，利于学习，使用广泛，主要适用于 Web 开发领域，但它丰富的图形库意味着它可以应用于数据的可视化。

六、Python 数据分析与可视化常用类库

1. Numpy

Numpy 软件包是 Python 生态系统中数据分析，机器学习和科学计算的主力军。它极大地简化了向量和矩阵的操作处理方式，它提供了许多高级的数值编程工具，如矩阵数据类型、矢量处理，以及精密的运算库，专为进行严格的数字处理而产生，多为很多大型金融公司及核心的科学计算组织使用，如 Lawrence Livermore、NASA 用其处理一些本来使用 C++、Fortran 或 Matlab 等所做的任务。

2. SciPy

SciPy 库依赖于 NumPy，它提供了便捷且快速的 N 维数组操作。SciPy 库的构建与 NumPy 数组一起工作，并提供了许多用户友好和高效的数字实践，可便捷地解决科学计算中的一些标准问题。

3. Pandas

Pandas 是基于 NumPy 的一种工具，该工具是为了解决数据分析任务而创建的。它纳入了大量库和一些标准的数据模型，提供了高效地操作大型数据集所需的工具，提供了大量能使我们快速便捷地处理数据的函数和方法。

3. Matplotlib

Matplotlib 是一个 Python 的 2D 绘图库，它以各种硬拷贝格式和跨平台的交互式环境生成出版质量级别的图形。它可与 Numpy 一起使用，提供一种有效的 MATLAB 开源替代方案；它也可以和图形工具包一起使用，让用户很轻松地将数据图形化；同时它还提供多样的输出格式。

4. Seaborn

Seaborn 是基于 Matplotlib 的图形可视化 Python 包。它提供了一种高度交互式界面，



便于用户能够做出各种有吸引力的统计图表。

5. Scikit-learn

基于 SciPy，开发者们针对不同的应用领域发展出的众多的分支版本被统一称为 Sci-kits，即 SciPy 工具包的意思。而在这些分支版本中，最有名，也是专门面向机器学习的一个就是 Scikit-learn。Scikit-learn 是 Python 语言中专门针对机器学习应用而发展起来的一款开源框架。作为专门面向机器学习的 Python 开源框架，内部实现了各种各样成熟的算法，容易安装和使用，样例丰富，而且教程和文档也非常详细，但它不支持深度学习和强化学习，也不支持图模型和序列预测，不支持 Python 之外的语言，不支持 PyPy，也不支持 GPU 加速。Scikit-learn 的基本功能，主要被分为六大部分：分类、回归、聚类、数据降维、模型选择和数据预处理。

此外，用于展现分析商用分析工具有 Style Intelligence、RapidMiner Radoop、Cognos、BO、Microsoft Power BI、Oracle、Microstrategy、QlikView、Tableau。

国内的有 BDP、国云数据（大数据魔镜）、思迈特、FineBI 等；数据仓库有 Teradata AsterData、EMC GreenPlum、HP Vertica 等；数据集市有 QlikView、Tableau、Style Intelligence 等。

七、大数据分析应用成果

1. 积极主动与预测需求

企业机构面临着越来越大的竞争压力，它们不仅需要获取客户，还要了解客户的需求，以便提升客户体验，并发展长久的关系。客户通过分享数据，降低数据使用的隐私级别，期望企业能够了解他们，形成相应的互动，并在所有的接触点提供无缝体验。

为此，企业需要识别客户的多个标识符（例如手机、电子邮件和地址），并将其整合为一个单独的客户 ID。由于客户越来越多地使用多个渠道与企业互动，为此需要整合传统数据源和数字数据源来理解客户的行为。此外，企业也需要提供情境相关的实时体验，这也是客户的期望。

2. 缓冲风险与减少欺诈

安全和欺诈分析旨在保护所有物理、财务和知识资产免受内部和外部威胁的滥用。高效的数据和分析能力将确保最佳的欺诈预防水平，提升整个企业机构的安全：威慑需要建立有效的机制，以便企业快速检测并预测欺诈活动，同时识别和跟踪肇事者。

将统计、网络、路径和大数据方法论用于带来警报的预测性欺诈倾向模型，将确保在被实时威胁检测流程触发后能够及时做出响应，并自动发出警报和做出相应的处理。数据管理以及高效和透明的欺诈事件报告机制将有助于改进欺诈风险管理流程。

此外，对整个企业的数据进行集成和关联可以提供统一的跨不同业务线、产品和交易的欺诈视图。多类型分析和数据基础可以提供更准确的欺诈趋势分析和预测，并预测未来的潜在操作方式，确定欺诈审计和调查中的漏洞。

3. 提供相关产品

产品是任何企业机构生存的基石，也通常是企业投入最大的领域。产品管理团队的作用是辨识推动创新、新功能和服务战略路线图的发展趋势。



通过对个人公布的想法和观点的第三方数据源进行有效整理,再进行相应分析,可以帮助企业在需求发生变化或开发新技术的时候保持竞争力,并能够加快对市场需求的预测,在需求产生之前提供相应产品。

4. 个性化与服务

公司在处理结构化数据方面仍然有些吃力,并需要快速应对通过数字技术进行客户交互所带来的不稳定性。要做出实时回应,并让客户感觉受到重视,只能通过先进的分析技术实现。大数据带来了基于客户个性进行互动的机会。这是通过理解客户的态度,并考虑实时位置等因素,从而在多渠道的服务环境中带来个性化关注实现的。

5. 优化与改善客户体验

运营管理不善可能会导致无数重大的问题,这包括面临损害客户体验,最终降低品牌忠诚度的重大风险。通过在流程设计和控制,以及在商品或服务生产中的业务运营优化中应用分析技术,可以提升满足客户期望的有效性和效率,并实现卓越的运营。

通过部署先进的分析技术,可以提高现场运营活动的生产力和效率,并能够根据业务和客户需求优化组织人力安排。数据和分析的最佳化使用可以带来端对端的视图,并能够对关键运营指标进行衡量,从而确保持续不断的改进。

任务四 应用 MySQL 数据库建立财务模型

MySQL 提供了丰富的函数,包括数学函数、字符串函数、日期和时间函数、条件函数、系统信息函数、加密函数、格式化函数等。掌握 MySQL 中的函数语法以及理解财务有关的常用函数,对我们以后在财务工作中进行数据处理有很大的帮助。

在 Excel 中,我们经常使用一些内置的函数来进行数据处理,那么在 MySQL 中有理数据的函数呢? MySQL 也提供了与数据处理相关的函数,如计数、最大值、最小值、分组和排序等。掌握这些函数,可以帮助我们提高工作效率。

一、MySQL 函数通用语法

鉴于需要较多的数据才能更好地体现出数据的分析与整理,在 MySQL 中,函数的使用语法如下:

select 函数名(字段名称) from 表名 where XXX = 'yyyf'

其中,视业务需求加 Where 条件。

二、财务整理与分析常用的函数

(一) 计数(count)

最重要的聚合函数之一是 count() 函数,它从表中的列返回记录数。可以用 count() 函数来获取总共销售过几笔,SQL 语句为 select count(product_name) from accounts_info; 当然, count() 函数里面的字段可以任意使用,这里我们只是统计销售过的交易数。

如果这时候我们要统计卖出多少种产品,会发现统计出来的产品有重复。那么怎么



去重呢？这里介绍一个经常与 count（计数）函数搭档使用的 distinct（）函数去重函数。参考代码为 select count（distinct（product _ name））from accounts _ info；结果为 10，因为将重复的“产品 A”去除了。

（二）求和（sum）

每当处理与数字有关的列时，必然要检查它们的总和。MySQL 中统计总和用 sum（）函数。大家可以先试着在自己的 Workbench 上写 SQL 语句，算出 accounts _ info 表中收入总和是多少。

“select sum（income）from ac. counts _ info sum；”语句中括号里面放的就是要统计的字段名，这样就能算出收入总和。

（三）最大（max）与最小（min）

在会计中，我们经常会碰到要统计最大值和最小值，例如，我们来统计一下 accounts _ info 表中的最大销售量和最小销售量。

MySQL 求最大用的函数是 max（），求最小用的函数是 min（）。大家可以先在自己的 Workbench 上练习，算出最大销售量和最小销售量。

首先，销售量的字段是“sale _ num”，所以很快就能写出“select max（sale _ num）from accounts _ info；”单击“运行”，得出销售最好成绩是 100 件；写出“select min（sale _ num）from accounts _ nlb；”单击“运行”，销售最差成绩是 12.97 件。

（四）分组（group by）和排序（order by）

很多时候我们要统计的不是一条条的记录，而是一个个组别的数据。例如，我们要统计每种产品的销售量，就像“产品 A”有两条记录，就应该用分组的形式。

先介绍分组的语法，分组通常用“group by”字段名“group by”通常是放在表名后面，如果有 where XXX = ‘yyy’，则放在 where 条件后面。以查询所有产品的销售量为例，大家可以先尝试在 Workbench 上写代码。

参考代码：

```
select sum（sale _ num）as sale _ sum, product _ name from accounts _ info group by product _ name；
```

分组 SQL 语句解析如图 1-3 所示。这时候就会发现每个产品的销售总和出现了，但是还不能一眼看出哪个销售量最高，所以还要进行排序。MySQL 的排序语法为“order by 字段名称”，“order by”需要放在表名之后，如果有 where 条件，要放在“where”之后。大家可以先在自己的 Workbench 上尝试写代码。

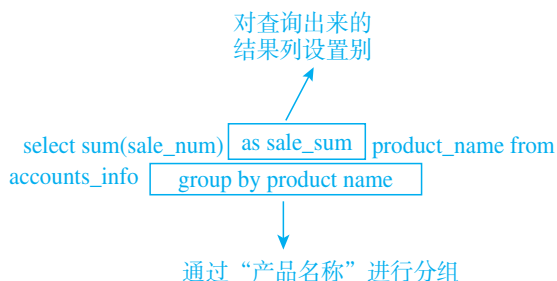


图 1-3

参考代码：

```
select sum (sale _ num) as sale _ sum, product _ name from accounts _ info group  
by product _ name order by sale _ sum;
```

现在排序效果有了，但是我们需要把最高的销售量排在第一位，那怎么办呢？排序有两种：升序和降序。默认的是升序。升序使用关键字“asc”，降序使用关键字“desc”，具体语法为 order by 字段名 asc/desc。这里我们用刚才的语句写一个降序的 SQL。

参考代码：

```
select sum (sale _ num) as sale _ sum, product _ name from accounts _ info group  
by product _ name order by sale _ sum desc;
```

(五) 绝对值 (abs)

在实际工作中，很多地方需要用到取绝对值，那么在 MySQL 中如何使用绝对值函数呢？

MySQL 绝对值函数为 abs (字段名)，这样就能对针对的字段取绝对值了。因为我们的数据没有负数，所以取销售量减去 100 作为尝试。先看看销售量减去 100 的结果是怎样的。

参考代码：

```
select sale _ num-100 as newSale from accounts _ info;
```

结果全部是负值。

现在来看绝对值怎么写。先尝试在本地的 Workbench 上写 SQL 语句。

参考代码：

```
select abs (sale _ num-100) as newSale from accounts _ info;
```

使用了 abs () 函数后，结果就都是正值了。

很多时候我们需要将数据进行小数点截取，不要小数，这时候通常有三种做法：

- (1) 向上取整，比如 84.67 变成 85。
- (2) 向下取整，比如 84.67 变成 84。
- (3) 进行四舍五入，比如 89.06 变成 89，89.50 变成 90。

这三种做法分别对应的是向上取整函数 ceil ()、向下取整函数 floor ()、四舍五入函数 round ()。

如果对 accounts _ info 表中的销售数量 sale _ num 进行向上取整，那么代码为：

```
select ceil (sale _ num) from accounts _ info;
```



如果对 accounts_info 表中的销售数量 sale_num 进行向下取整，那么代码为：

```
select floor (sale_num) from accounts_info;
```

如果对 accounts_info 表中的销售数量 sale_num 进行四舍五入取整，那么代码为

```
select round (sale_num) from accounts_info;
```

这里我们着重介绍一下四舍五入 round () 方法，这个方法不像向上、向下取整的结果只能是整数，而是可以用于取指定位数的小数。例如，我们在取销售数量的时候，结果总是有很多位小数点，如果只想取两位小数点，那么该怎么写呢？

保留指定小数位并进行四舍五入的语法为：round (字段名，保留的小数位)，大家可以在本地的 Workbench 上尝试写 SQL 语句。

参考代码：

```
select round (sale_num, 1) from accounts_info;
```

课后练习

一、简答题

- (1) 什么是大数据？大数据的特点是什么？
- (2) 大数据分析内容是什么？
- (3) 数据分析与可视化常用工具有哪些？

二、实践操作

python 基础 (<http://it.tsgcjy.cn/pythonnew/index.html>)