



# 商务智能

薛云 / 主编

郭彦丽 谢桂袖 / 副主编

扫一扫看本书资源



01

“教师+企业工程师”  
构成权威的编写团队

02

遵循理论够用和实践  
充分的原则，案例丰富

03

教学资源丰富，提供  
全套教学解决方案



# 商务智能

薛云 / 主编

郭彦丽 谢桂袖 / 副主编

人民邮电出版社

北京

## 图书在版编目 (C I P) 数据

商务智能 / 薛云主编. -- 北京 : 人民邮电出版社,  
2019.3  
ISBN 978-7-115-43976-5

I. ①商… II. ①薛… III. ①数据处理 IV.  
①TP274

中国版本图书馆CIP数据核字(2018)第270004号

## 内 容 提 要

商务智能是近年来企业信息化的热点,有着广阔的应用前景。本书以“认识—实施—案例—工具—发展”为主线,介绍了商务智能从理论到实践的全过程。全书一共分为7章,内容包括:第1章“认识商务智能”,解释了商务智能的相关概念及理论基础;第2章“实施商务智能”,讲述了商务智能项目的实施过程;第3章“分析商务智能案例”,展示了商务智能实际案例的实施细节、成果及价值所在;第4章至第6章讲解了商务智能开发工具的应用,即应用微策略软件进行数据报表的分析与展现,并通过案例讲解了数据挖掘方法;第7章介绍了商务智能的应用及未来发展趋势。

本书内容由浅入深、案例丰富、实践性强,可引导读者掌握商务智能的基础知识、熟悉商务智能的应用场景,进而让读者能够进行实际操作。本书可以作为高等院校信息管理、计算机应用、电子商务等相关专业学生学习商务智能理论与实践的教材,也可以作为从事商务智能相关工作的人员的参考资料,还可以作为MBA技术管理课程的教材。

- ◆ 主 编 薛 云
- 副 主 编 郭彦丽 谢桂袖
- 责任编辑 朱海昀
- 责任印制 马振武
- ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号  
邮编 100164 电子邮件 315@ptpress.com.cn  
网址 <http://www.ptpress.com.cn>  
山东百润本色印刷有限公司印刷
- ◆ 开本: 787×1092 1/16  
印张: 14.25 2019年3月第1版  
字数: 320千字 2019年3月山东第1次印刷

定价: 49.80 元

读者服务热线: (010)81055256 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147 号

# 前言

## PREFACE

商务智能 (Business Intelligence, BI) 是利用数据存储、数据集成、数据分析技术, 把海量的业务数据转化成有价值的信息, 辅助管理者进行决策的一套完整的解决方案。目前, 各行各业对商务智能人才的需求也在不断增加。很多大学的计算机学院、软件学院和商学院等都开设了商务智能的相关课程, 培养商务智能人才, 以满足市场的需求。因此, 编者从国内商务智能人才培养的需求出发, 考虑到国内外相关著作中存在的不足之处, 结合北京汇智协同教育科技有限公司推出的“CEAC 职业技能培训测评——商务智能课程”撰写本书, 并开发了相应的视频教学资源。希望本书能够符合市场需求, 也希望读者通过学习本书能够对商务智能有更深入的了解。

全书以“认识—实施—案例—工具—发展”为主线, 讲述了商务智能从理论到实际操作的全过程。“认识”部分解释了商务智能的相关概念及理论基础, 让读者充分了解商务智能; “实施”部分讲述了商务智能项目的实施过程, 为商务智能的实际操作打牢知识基础; “案例”部分展示了企业真实商务智能案例的实施细节、成果及价值, 让读者亲身体验商务智能带给企业决策者的洞察力; “工具”部分应用微策略 (MicroStrategy, MSTR) 软件进行实践, 让读者体验商务智能工具的便捷和强大; “发展”部分介绍了商务智能的应用及未来发展趋势, 引导读者追随技术发展的方向。

本书的特色如下。

(1) 结构合理: 按照理论够用和实践充分的原则进行综合构思和组织内容。

(2) 资源丰富: 考虑到读者的学习需求, 本书配备了丰富的教学资源。读者可以登录 CEAC 职业技能培训测评学习网站, 选择“在线课程”专区的商务智能课程进行学习, 并获取相关参考资料。读者也可以关注微信公众号 (LearningBI), 获取相关学习资料, 参与相关的互动。

(3) 实践充分: 本书提供了 Desktop 工具的下载链接、数据文件和微课视频。读者可以在人邮教育社区 (www.ryjiaoyu.com) 下载获取。

(4) 案例具体: 书中的案例从实际的企业背景和商业需求出发, 进行了功能的展示, 使读者可以充分了解商务智能的实践效果, 体现了本书的实用性。

本书由薛云担任主编, 郭彦丽、谢桂袖担任副主编, 参与编写的还有胡佳伟、林曼、吴晟和刘秋霞。各章具体编写分工如下: 薛云负责第 1、2 章的编写; 胡佳伟、林曼和吴晟负责第 3 章的编写, 薛云和刘秋霞负责第 4、5、6 章的编写; 郭彦丽、谢桂袖负责第 7 章的编写。

由于编者水平所限，书中难免存在疏漏之处，敬请广大读者批评指正，以期在今后的修订中进一步完善。

特别感谢北京汇智协同教育科技有限公司对本书编写进行非常具体的规划和设计，并跟踪和把控本书编写的整个过程。与此同时，感谢德昂信息技术（北京）有限公司在师资、教学资源、教学案例等方面给予的大力支持。编者在本书编写过程中参考了有关书籍和资料，在此向相关作者表示衷心感谢！

编者

2018年10月

# CONTENTS 目录

## 第 1 章

### 认识商务智能..... 1

#### 1.1 什么是商务智能..... 2

1.1.1 商务智能的概念..... 2

1.1.2 商务智能的特点..... 2

1.1.3 商务智能的发展历史..... 3

1.1.4 商务智能的技术实现路径..... 4

1.1.5 商务智能的架构模型..... 5

1.1.6 商务智能的用户..... 6

1.1.7 商务智能与大数据..... 6

#### 1.2 什么是数据可视化..... 8

1.2.1 数据可视化的概念..... 8

1.2.2 商务智能与数据可视化的关系..... 8

1.2.3 常见图表概览..... 9

#### 1.3 什么是数据仓库..... 13

1.3.1 数据仓库的定义..... 13

1.3.2 数据仓库的特点..... 14

1.3.3 数据仓库的建模..... 14

1.3.4 数据集市的定义..... 17

1.3.5 数据仓库的体系结构..... 18

1.3.6 数据仓库的数据及组织..... 19

#### 1.4 什么是 ETL..... 19

1.4.1 数据的抽取..... 20

1.4.2 数据的清洗..... 20

1.4.3 数据的转换..... 21

#### 1.5 什么是数据挖掘..... 21

1.5.1 数据挖掘的定义..... 21

1.5.2 数据挖掘的功能..... 22

1.5.3 数据挖掘的对象..... 23

1.5.4 数据挖掘的步骤..... 24

1.5.5 数据挖掘在商务智能中的应用..... 25

课后习题..... 26

## 第 2 章

### 实施商务智能..... 27

#### 2.1 商务智能的准备与规划..... 28

2.1.1 主要内容..... 28

2.1.2 阶段产物..... 30

#### 2.2 商务智能的需求分析..... 31

2.2.1 主要内容..... 31

2.2.2 阶段产物..... 37

#### 2.3 商务智能系统的设计..... 37

2.3.1 主要内容..... 38

2.3.2 阶段产物..... 40

#### 2.4 商务智能系统的实现..... 40

2.4.1 主要内容..... 40

2.4.2 阶段产物..... 42

#### 2.5 商务智能系统的上线及维护..... 43

2.5.1 主要内容..... 43

2.5.2 阶段产物..... 43

课后习题..... 44

### 第 3 章

#### 分析商务智能案例 ..... 45

##### 3.1 数据可视化趣味案例 ..... 46

3.1.1 《唐诗三百首》 ..... 46

3.1.2 周杰伦歌曲 ..... 49

3.1.3 民间快餐 ..... 53

3.1.4 “中国有嘻哈” 节目 ..... 57

##### 3.2 财经行业案例 ..... 61

3.2.1 案例背景 ..... 61

3.2.2 项目实践 ..... 62

3.2.3 项目展示 ..... 67

3.2.4 项目价值 ..... 70

##### 3.3 快速消费品行业案例 ..... 71

3.3.1 案例背景 ..... 71

3.3.2 项目实践 ..... 72

3.3.3 项目展示 ..... 75

3.3.4 项目价值 ..... 79

##### 3.4 餐饮行业案例 ..... 79

3.4.1 案例背景 ..... 79

3.4.2 项目实践 ..... 80

3.4.3 项目展示 ..... 82

3.4.4 项目价值 ..... 84

##### 3.5 医药行业案例 ..... 85

3.5.1 案例背景 ..... 85

3.5.2 项目实践 ..... 86

3.5.3 项目展示 ..... 87

3.5.4 项目价值 ..... 90

#### 课后习题 ..... 91

### 第 4 章

#### 商务智能开发工具

#### ——MicroStrategy

#### Developer ..... 92

##### 4.1 开发工具概述 ..... 93

4.1.1 基础知识 ..... 93

4.1.2 开发方式 ..... 94

4.1.3 MicroStrategy Developer

简介 ..... 96

##### 4.2 项目创建 ..... 97

4.2.1 创建元数据库 ..... 97

4.2.2 连接项目源 ..... 100

4.2.3 创建项目 ..... 101

##### 4.3 数据定义 ..... 103

4.3.1 定义数据库实例 ..... 103

4.3.2 选择数据仓库表 ..... 105

4.3.3 定义事实 ..... 106

4.3.4 定义实体 ..... 112

4.3.5 更新框架 ..... 121

##### 4.4 报表创建 ..... 122

4.4.1 创建度量 ..... 122

4.4.2 创建报表 ..... 124

##### 4.5 网页浏览 ..... 126

4.5.1 设置智能服务器 ..... 126

4.5.2 浏览报表 ..... 131

4.5.3 以服务器方式连接元数据库 ..... 131

#### 课后习题 ..... 132

### 第 5 章

#### 商务智能开发工具——

#### MicroStrategy

#### Desktop-Dossiers ..... 133

##### 5.1 开发工具概述 ..... 134

##### 5.2 数据连接 ..... 134

##### 5.3 数据清洗 ..... 138

##### 5.4 数据可视化 ..... 144

5.4.1 收益分析 ..... 144

5.4.2 收益报表细化 ..... 151

5.4.3 销售业绩和利润率 ..... 163

5.4.4 关键指标 ..... 166

##### 5.5 仪表盘美化 ..... 171

5.5.1 添加图像 ..... 171

5.5.2 修改可视化效果 ..... 172

##### 5.6 仪表盘分享 ..... 175

5.6.1 导出成 PDF 格式 ..... 175

5.6.2 Desktop 用户 ..... 175

#### 课后习题 ..... 176

### 第 6 章

#### 商务智能开发工具

#### ——MicroStrategy Data Mining

#### Services ..... 177

##### 6.1 开发工具概述 ..... 178

6.1.1	MicroStrategy 的数据挖掘功能	178
6.1.2	MicroStrategy 的数据挖掘方式	179
6.1.3	PMML 概要	179
6.1.4	应用数据挖掘服务的工作流程	180
<b>6.2</b>	<b>线性回归和季节回归</b>	<b>180</b>
6.2.1	目标	180
6.2.2	数据准备	180
6.2.3	数据挖掘步骤	181
<b>6.3</b>	<b>逻辑回归</b>	<b>189</b>
6.3.1	目标	189
6.3.2	数据准备	189
6.3.3	数据挖掘步骤	190
<b>6.4</b>	<b>决策树分析</b>	<b>193</b>
6.4.1	目标	193
6.4.2	数据准备	193
6.4.3	数据挖掘步骤	194
<b>6.5</b>	<b>关联规则</b>	<b>197</b>
6.5.1	目标	197

6.5.2	数据准备	197
6.5.3	数据挖掘步骤	197

<b>6.6</b>	<b>聚类分析</b>	<b>202</b>
6.6.1	目标	202
6.6.2	数据准备	202
6.6.3	数据挖掘步骤	203
	课后习题	205

## 第 7 章

	展望商务智能的未来	207
<b>7.1</b>	<b>商务智能的应用</b>	<b>208</b>
7.1.1	商务智能的应用范围	208
7.1.2	商务智能的应用价值	215
<b>7.2</b>	<b>商务智能的发展趋势</b>	<b>217</b>
	课后习题	218
	参考文献	219



# 第 1 章

## 认识商务智能

据统计，在如今信息爆炸的时代，全球每天大约产生 15PB（1PB=1024TB）数据，且数据量正在以每年 1.3 倍的速度迅速扩增。然而在海量的数据中，真正能够被企业用于分析和运用的部分不足 10%。国际数据企业的调查表明，企业 37% 的业务决策主要依赖于“直觉”或“本能”。面对日益激烈的市场竞争，企业越来越需要提高决策的准确度和灵敏度。显然，依赖于“直觉”和“本能”的决策模式已经不符合时代的发展。企业现有的海量数据是隐藏着巨大商业价值的资产，如何从这些海量数据中挖掘出企业决策者所需的信息已经越来越重要。在此背景下，商务智能（Business Intelligence, BI）应运而生，而且得到相关业界的日益重视。

本章将从商务智能概念，数据可视化，数据仓库，数据的抽取、转换、加载，数据挖掘几个方面认识和解读商务智能。

## 【学习目标】

1. 了解商务智能的发展历史，理解商务智能的基础知识，掌握商务智能的架构模型。
2. 了解数据可视化的概念，理解商务智能与数据可视化的关系，掌握常见图表的应用场景。
3. 理解数据仓库的基础知识，掌握数据仓库的建模及体系结构。
4. 理解数据抽取、转换、加载的含义。
5. 了解数据挖掘的定义及在商务智能中的应用，理解数据挖掘的功能和对象，掌握数据挖掘的步骤。

## 1.1 什么是商务智能

### 1.1.1 商务智能的概念

从 20 世纪 90 年代开始到目前，商务智能越来越受到企业界的关注，其概念最早由美国加特纳企业（Gartner 是全球最具权威的 IT 研究与顾问咨询企业之一）于 1996 年提出。

加特纳企业将商务智能定义为：商务智能描述了一系列的概念和方法，通过应用基于事实的支持系统来辅助商业决策的制订，商务智能技术提供使企业迅速分析数据的技术和方法，包括收集、管理和分析数据，将这些数据转化为有用的信息，然后分发到企业各处。

国内研究商务智能的著名学者王茁，给出的商务智能的概念为：商务智能是企业利用现代信息技术收集、管理和分析结构化与非结构化的商务信息，创造和累计商务知识、见解，进而改善商务决策水平、采取有效的商务行动、完善各种商务流程、提升各方面商务绩效、增强综合竞争力的智慧和能力。

综上，对商务智能概念的理解可以分别从信息技术和管理科学的角度进行。从信息技术的角度理解，商务智能是一种满足企业决策需要的解决方案，即从来自不同的企业运作系统的数据中提取有用的数据并进行清理，以保证数据的正确性，然后经过抽取、转换和加载，再合并到一个企业级的数据仓库里，从而得到企业数据的全局视图，并在此基础上利用合适的查询和分析工具、数据挖掘工具等对其进行分析和处理，最后为管理者的决策过程提供支持。从管理科学的角度理解，商务智能是指对商业信息的搜集、管理和分析的过程，目的是使企业的各级决策者获得知识、提升洞察力，促使他们做出对企业更有利的决策。

### 1.1.2 商务智能的特点

商务智能是融合了先进信息技术与创新管理理念的结合体，我们通过了解商务智能的特点可以更好地理解商务智能的内涵。商务智能的特点如下。

#### 1. 综合性的开放系统

商务智能是企业面向内外部环境，同外界环境保持动态互连的开放系统。

#### 2. 具有强大的数据分析处理与展示功能

商务智能集成了在线分析处理、数据挖掘等多项数据分析技术。

#### 3. 注重在系统的海量数据和信息中发现知识

企业为了在竞争中取得优势地位，必须通过商务智能技术识别和应用隐藏在所收集的数据中的知识。

#### 4. 综合了多项技术的应用

商务智能所采用的技术并不是新的技术，而是已有的数据仓库、在线分析处理、数据挖掘等技术的综合。

#### 5. 服务于企业战略

商务智能对企业的内外部数据进行分析，支持企业战略管理。

#### 6. 有助于提升企业绩效

商务智能必须要促进企业某一方面业务顺利开展，提升业绩。

#### 7. 用户具有多样性

商务智能的用户包括企业一线的业务人员、各级管理者，甚至外部的顾客和商业合作伙伴，每一层用户拥有不同的使用权限，不过商务智能最终服务于各类企业决策者。

### 1.1.3 商务智能的发展历史

从最初的事务处理系统（Transaction Processing System, TPS），到高层管理信息系统（Executive Information System, EIS）、管理信息系统（Management Information System, MIS）和决策支持系统（Decision Support System, DSS）等，再到今天的企业商务智能，就是商务智能的发展历程。

#### 1. TPS

TPS 是企业信息化进程中首要进行的任务之一。TPS 是进行日常业务处理、记录、汇总、综合、分类，并为组织的操作层次服务的基本商务系统。TPS 可以帮助组织降低业务成本，提高信息准确度，提升业务服务水平，其在企业中主要表现为 4 种系统：市场营销系统，生产制造系统，财务系统会计系统，人力资源系统。TPS 向 EIS、MIS、DSS 和商务智能系统提供了所要的基础数据，是它们的基础。

#### 2. EIS

EIS 是服务于组织高层经理的一类特殊的信息系统，能够使经理们更快地得到更广泛的信息。EIS 首先是一个“组织状况报导系统”，能够迅速、方便、直观地用图形提供综合信息，并可以预警与控制“成功关键因素”遇到的问题，能有选择地向管理人员和执行人员提供关于业务状况的信息。EIS 虽然能提供关于商业活动情况的一些信息，但若要对商业活动面临的问题进一步分析，还要借助于另一些分析工具或由专业人员来实现。

#### 3. MIS

MIS 由人和计算机网络集成，能提供企业管理所需信息，以支持企业的生产经营和决策的人机系统，主要功能包括经营管理、资产管理、生产管理、行政管理和系统维护等。MIS 是 EIS 的进一步发展，其应用范围比 EIS 更为广泛，能够帮助管理人员了解日常业务，并进行高效的控制、组织、计划。

#### 4. DSS

DSS 是辅助决策者通过数据、模型和知识，以人机交互方式进行半结构化或非结构化决策的计算机应用系统。它为决策者提供分析问题、建立模型、模拟决策过程和方案的环境，调用各种信息资源和分析工具，帮助决策者提高决策水平和质量。它是 MIS 向更高级发展而产生的先进信息管理系统。

DSS 比 MIS 更为灵活，它允许决策者查询存储于关系数据库中的任何问题，甚至储存于不同计算机系统或网络中的数据库中的有关数据，并以多样化的格式提交给决策者及其他信息系统。

## 5. 商务智能

随着互联网的快速发展，在 DSS 基础上发展商务智能成为必然。因为在统一的平台上，企业能向组织内外的人员（包括员工、供货商、合作伙伴、客户）方便地发送信息；而且，随着基于互联网的各种信息系统在企业中的应用，企业将收集越来越多的关于客户、产品及销售情况在内的各种信息，这些信息能帮助企业更好地预测和把握未来。在 DSS 基础上进一步发展起来的商务智能系统能够向用户提供更为复杂的商业信息，可以更为方便地定制各种报表和图表的格式，能够向行政管理人员、技术人员和普通员工提供个性化的多维信息，使分析处理信息的能力和信息的利用率大为提高。例如，用户想了解销售情况时，可以通过商务智能系统得到按产品、地区、客户分类的网上销售和正常柜台销售的多种分析报告，在此基础上，可以进一步解决企业决策时需要了解的各种问题，并帮助企业更快、更好地制订和做出决策。

另外，随着企业信息技术的升级，如今在企业界，数据资产的观念正在进入企业的资源计划（Enterprise Resource Planning, ERP）系统中，而把数据转换为资产的方法和技术也正在成为企业投资 IT 领域的热点。目前大部分大中规模的企业都是信息丰富的组织，而一个信息丰富的组织的绩效不仅仅依赖于产品、服务或地点等因素，更重要的是依赖知识。从数据转换为信息，再从信息转换为知识，并不是一个简单的过程。商务智能的本质正是把数据转化为知识，其致力于知识发现和挖掘，使企业的数据资产能带来明显的经济效益，减少不确定性因素的影响，使企业取得新的竞争优势。

### 1.1.4 商务智能的技术实现路径

商务智能的技术实现路径参见图 1-1。



图 1-1 商务智能的技术实现路径

首先，数据是起点，也是基础。ERP 系统中已经积累了大量数据，但是，这些数据却是按照单据与流程的需要而存储的，对于管理者来说，就显得有些杂乱无章。他们期望看到的数据是简单且直观的，所以，分析人员需要将数据进行抽取、清洗，将之转换为有价值的信息。

其次，信息转变为知识，而知识管理的一个重要工作就是将某个特定人脑中的经验，变为可复制。在这里，这个过程就是建模的过程。将不同分析主题的分析视角（维度）与分析内容（度量）固化下来，让大家知道原来可以从这些角度来分析这么多指标。

最后，知识辅助决策。决策不是少数高管的专利。管理学中有一个著名的“木桶理论”，就是说，一个木桶能装多少水，并不取决于木桶最长的木板，而取决于最短的那根。而对于企业管理来说，不管董事长、总经理的决策水平多高，相关决策是否能发挥应有作用的关键在于各级管理人员都能理解且执行到位。所以，通过一个数据分析平台的建立，让所有的管理者都看到统一的数据（当然是有限控制的），都能基于数据去决策与执行，才能真正提升企业的整体决策水平。

因此，商务智能的技术实现路径，总结下来有3点：数据获取、建模与平台化展现。企业的商务智能离不开对数据的获取、使用和管理这3个过程，如图1-2所示。

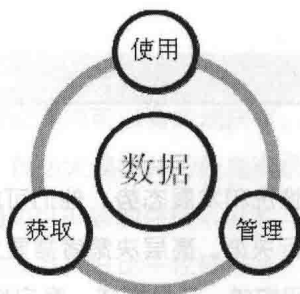


图 1-2 企业数据的应用过程

### 1.1.5 商务智能的架构模型

商务智能系统是一系列的概念、方法和过程的集合体，通过这些概念、方法和过程来获取和分析数据，提取有用的信息，更好地帮助决策。企业级商务智能系统可以看成是一种解决方案，它能够帮助企业通过现有的数据资源，获取和分析信息，帮助企业管理者做出最优决策。将企业级商务智能系统的层次架构划为数据层、技术层、分析层、展示层、决策层这5个层次，如图1-3所示。

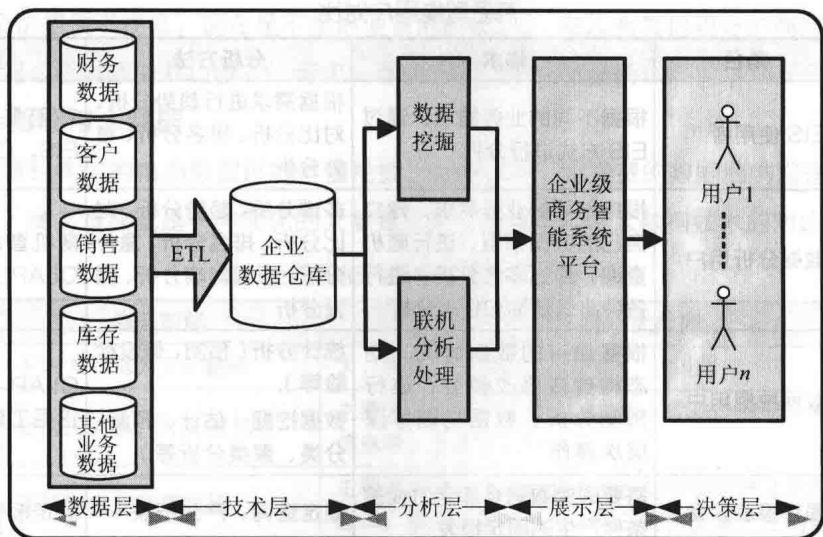


图 1-3 企业商务智能模型

(1) 数据层，确保从企业的 ERP 系统、客户关系管理系统 (Customer Relationship Management, CRM)、供应链管理系统 (Supply Chain Management, SCM) 等系统收集到真实的、有效的数据。

(2) 技术层，商务智能系统通过 ETL (Extract-Transform-Load, 数据抽取、转换、加载) 将数据层的原始数据集成到数据仓库中，针对不同部门的异构数据进行整合，以待进一步的分析处理。

(3) 分析层，系统需建立良好的模型库、知识库、方法库，从数据仓库中分析和挖掘出有价值的信息，转化为用户能理解的知识，充分展现企业级数据的智能分析功能。

(4) 展示层, 企业可通过查询报表, 制订关键绩效指标, 进行绩效管理等工作。

(5) 决策层, 用户运用系统提供分析结果, 将战略决策用于指导具体的行动, 体现了商务智能的价值。

### 1.1.6 商务智能的用户

#### 1. 高层决策者

高层决策者需要了解业务的总体情况和发展态势。他们可能使用系统提供的分析工具自己发现问题, 但更主要的是利用分析结果进行决策。高层决策者需要通晓业务的具体状态和发展趋势, 包括业务的状态和构成(机构构成、时间构成、产品构成、客户构成等), 以及各个指标的发展趋势和预测。

#### 2. 数据分析专家

数据分析专家需要更加深入地从数据仓库中发现问题和市场机会及风险, 需要及时把发现的结果报告给高层决策者。

#### 3. 中下级经理和业务人员

中下级经理和业务人员, 通常仅仅关心与各自工作相关的内容, 注重报表和固定的数据查询。

商务智能的用户类型、角色、需求、分析方法及所需的前端工具如表 1-1 所示。

表 1-1 商务智能用户对比

用户类型	角色	需求	分析方法	前端工具
高层决策者	EIS 使用者	根据不同的业务需求, 通过 EIS 方式进行分析	根据需求进行趋势分析、对比分析、排名分析、意外分析	利用各种软件开发的 EIS、OLAP 分析工具
数据分析专家	数据分析用户	根据不同的业务需求, 建立自己的数据模型, 进行随机查询; 通过多维分析, 进行各种高级查询和报表分析	多维分析、趋势分析、对比分析、排名分析、意外分析、原因影响分析、假设分析	随机查询及报表工具、OLAP 分析工具
	数据挖掘用户	根据现有的数据情况, 动态构建或修改模型, 进行预测分析、数据挖掘等深层次操作	统计分析(预测、假设检验等), 数据挖掘(估计、预测、分类、聚类分析等)	OLAP 分析工具、数据挖掘工具
中下级别经理和业务人员	固定报表读者	需要阅读数据仓库定时或按条件产生的固定报表	固定查询、产生报表	固定报表工具
	信息浏览者	根据不同的业务需求, 通过建立简单的查询进行分析, 产生动态报表	自查询、动态报表	即席查询及报表工具

### 1.1.7 商务智能与大数据

#### 1. 大数据的概念

研究机构 Gartner 给出的大数据的定义是: 大数据是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力来适应海量、高增长率和多样化的信息资产。

麦肯锡全球研究所给出的大数据的定义是：一种规模大到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合，具有海量的数据规模、快速的数据流转、多样的数据类型和价值密度低四大特征。

## 2. 大数据的相关技术

大数据需要特殊的技术，在合理时间内可以有效地获取、管理、处理大量的数据，并整理成为有价值的信息。适用于大数据的技术，包括大规模并行处理数据库、数据挖掘技术、分布式文件系统、分布式数据库、云计算平台、互联网和可扩展的存储系统等。

从技术上看，大数据与云计算的关系就好比一枚硬币的正反面一样密不可分。大数据必然无法用单台的计算机进行处理，必须采用分布式架构。它的特色在于对海量数据进行分布式数据挖掘。但它必须依托云计算的分布式处理、分布式数据库和云存储、虚拟化技术。

## 3. 大数据的应用

大数据到底改变了什么？它其实就是通过海量的数据处理，让决策过程变得更轻松、更靠谱。比如，在网上购物时，它首先会告诉你某个商品有多少人评价，评价越多，也就意味着越多人买；当单击进去之后，就可以看到好评度的具体数据。还可以提供一个价格趋势信息，展示历史价位走势，告诉用户其他电商平台是不是更便宜。除此之外，还可提供其他同类或同价位商品的推荐与排名等。有了这些信息，用户就掌握了做购买决策时的几个关键信息：“这款商品好不好卖？”“这款商品口碑好不好？”“这个价格便不便宜，现在是不是最便宜的时候？”“有没有其他更好的商品？”“这个时候，是否将这个商品放进购物篮并付款？”可见，做决定将变得非常轻松。

## 4. 商务智能与大数据

大数据是将结构化的精确数据进行价值挖掘，化繁为简；将海量的数据归纳整理为几个简单的指标。简而言之，商务智能就是一种简化的大数据工具。商务智能与大数据技术的对比如表 1-2 所示。

表 1-2

商务智能与大数据技术的对比

内容	商务智能	大数据
信息量	不太大，常为 TB 量级	大，常为 PB 量级
信息特征	结构化信息，非实时信息	主要为非结构化信息，如文本、图形、音频、视频、遥感和遥测信息等，大多是实时信息
信息来源	主要为企业交易数据	主要是社会日常运作和各种服务中实时产生的数字数据，如社交媒体、移动电话和短信、电子商务交易产生的数据等
涉及技术	数据仓库、数据挖掘技术等	大规模并行处理数据库、数据挖掘技术、分布式文件系统、分布式数据库、云计算平台、互联网和可扩展的存储系统等
关联关系	商务智能是达成业务管理的应用工具。没有商务智能，大数据就没有了转化为价值的工具，没办法变成决策的依据	大数据是基础，没有大数据，商务智能就失去了存在的基础。同样，商务智能未来的发展方向是“极速、实时、智能”。要能够快速、实时地看到通过大数据产生的库存、账单等信息，还要建立业务模型和数据模型，让大数据按照业务逻辑展现
数据来源	结构化数据	大数据能够基于商务智能工具对大容量数据和非结构化数据进行处理；与传统基于事务的数据仓库系统相比较，大数据分析不仅关注结构化的历史数据，还倾向对非结构化海量数据进行分析；大数据是对商务智能的一个完美补充

续表

内容	商务智能	大数据
精确性要求	高	不高
因果与关联	重因果性分析	重关联性分析
效益	通过数据分析提高运营能力	将更有利于深入了解业务运转和与客户的互动

## 1.2 什么是数据可视化

### 1.2.1 数据可视化的概念

数据可视化旨在借助图形化手段，来清晰、有效地传达信息。数据可视化为了达到上述目的，需在数据满足分析决策需要和数据展现形式上同时予以考虑，通过直观地传达关键数据与特征，实现对于相当分数而又复杂的数据集的深入洞察。然而，现实中，设计人员往往并不能很好地把握设计与功能之间的平衡，从而创造出华而不实的数据可视化形式，无法达到其主要目的，也就是传达与沟通不畅。

数据可视化与信息图形、信息可视化、科学可视化以及统计图形密切相关。当前，在研究、教学和开发领域，数据可视化是一个极为活跃而又重要的领域。

### 1.2.2 商务智能与数据可视化的关系

商务智能的目标是将商业和企业运维中收集的数据转化为知识，辅助决策者做出明智的业务经营决策。数据包括来自业务系统的订单、库存、交易账目、客户和供应商等，以及其他外部环境中的各种数据。从技术层面上看，商务智能是数据仓库、联机分析处理工具和数据挖掘等技术的综合应用，其目的是使各级决策者获得知识、提升洞察力。为了使分析后的数据直观、简练地呈现在用户面前，则需要采用一定的形式表示和发布出来，此时便需要应用数据可视化技术。

商务智能中的数据可视化，是以商业报表、关键绩效指标、图形等易为人们所辨识的方式将原始数据间的复杂关系、潜在信息以及发展趋势，通过可视化展现平台，以易于访问和交互的方式来揭示数据的价值，从而提升决策人员的业务过程洞察力。目前，多数商务智能软件企业已提供了基于 Web 应用的展现服务，来扩展商务智能的信息发布范围。另外，随着移动应用的普及和移动办公的需求日益强烈，部分主流商务智能软件企业也有了移动端平台展现可视化开发的业务，例如微策略（MicroStrategy）。下面介绍两个在商务智能数据可视化中常用的两种展现形式，即仪表盘和平衡计分卡。

#### 1. 仪表盘

仪表盘（Dashboard）在商务智能分析中起着很重要的直观展示数据与支持决策的作用。由于仪表盘是一个商务智能综合分析展示的平台，所以一个界面上可能会有多个组件和图表，但这并不意味着这些数据和图表组件是随意堆砌的。商务智能仪表盘的展示大多有一个特定的主题或分类。大致归纳为 3 种：运营仪表盘、策略仪表盘和战略仪表盘。根据种类的不同，仪表盘的设计及展示也大不相同。

商务智能仪表盘也可以被称为管理驾驶舱，无论是管理决策者，还是企业业务流程中的普通员工，



都可以利用它来展示聚合分析的结果，让决策更加快速、准确，进而更快地推进业务流程，最终提高工作效率。

## 2. 平衡计分卡

平衡计分卡(Balanced Score Card)是商务智能分析中另一个主要应用数据可视化技术的部分。1992年,Robert Kaplan及David Norton在《Harvard Business Review》(哈佛商业评论)中发表题为“The Balanced Score Card: Measures That Drive Performance”的文章(平衡计分卡:驱动绩效的度量)。这是第一次提出平衡计分卡的概念。

平衡计分卡主要是通过图、卡、表来实现战略的规划。此工具可以将企业的策略,透过财务、内部业务流程、学习与成长、客户这4个方面来审视,如图1-4所示。每一方面都包括了策略目标、行动计划及衡量指标这三大部分。所谓“平衡”,是从3个角度来观察:一是外部及内部间的平衡,外部强调财务方面及客户方面,内部则强调内部流程及学习与成长方面;二是财务及非财务方面衡量的平衡;三是领先指标及落后指标的平衡等。

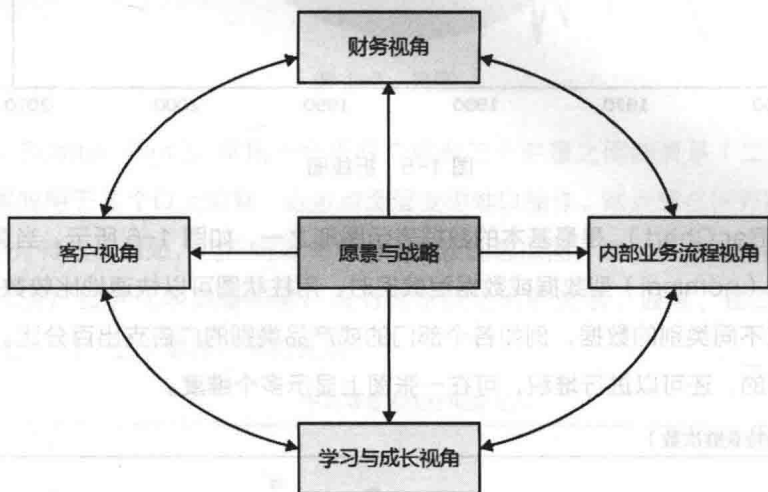


图 1-4 平衡计分卡视角图

### 1.2.3 常见图表概览

通常业务分析系统的终端用户在面对特定目标时,并不很确定应该使用何种类型图表。某些图表能够更好地解答特定的问题。下面简单介绍业务分析系统中常见的图表,并分别说明它们能够更好地解决或分析什么问题。有人觉得,基本图表太简单、太原始,不高端、不大气,因此追求更复杂的图表。这是对图表的一种误解。其实越简单的图表,越容易理解。快速、简洁地表现数据是“数据可视化”的最重要目的和最高的追求。

#### 1. 基本图表

下面介绍数据可视化中最常用的基本图表。

(1) 折线图(Line Chart),是最常用于描述时间序列数据的图表,如图1-5所示。折线图表示两种变量之间的关系,常用于追踪在一个时间段内的趋势或变化(通常将时间设为X轴)。折线图将