普通高等教育经管类专业系列教材

企业经营与财务智能分析可视化

(Excel、PowerBI、FineBI)

汪　刚　金春华　主　编
刘　畅　舒　艳　李　洁　副主编

清华大学出版社
北　京

内 容 简 介

本书详细介绍了企业经营与财务智能分析可视化的相关知识和具体应用。全书内容分为基础部分和工具应用部分：基础部分主要介绍商业智能基础、数据分析与可视化基础等理论知识；工具应用部分介绍 Excel、PowerBI、FineBI 等商业智能数据分析工具，其中重点介绍 PowerBI 的应用。

本书共有 12 章，每章按照“学习目标、引导案例、内容讲解、本章小结、本章习题”五大板块设计。

本书作为入门级教材，主要面向普通高等院校(包括应用型本科、高职)经管类相关专业的学生，以及对数据分析感兴趣的企事业单位人员，旨在使读者了解商业智能及数据分析可视化的基本原理，熟悉数据分析的步骤及思路，既能熟练使用 Excel 进行数据分析，又能掌握 PowerBI、FineBI 等常见自助式 BI 工具的一般应用。

本书配套了教案、教学课件、教学大纲、案例数据和微课等教学资源，方便教师教学和学生自学。

本书封面贴有清华大学出版社防伪标签，无标签者不得销售。
版权所有，侵权必究。举报：010-62782989，beiqinquan@tup.tsinghua.edu.cn。

图书在版编目(CIP)数据

企业经营与财务智能分析可视化：Excel、PowerBI、FineBI / 汪刚，金春华主编. —北京：清华大学出版社，2022.3

普通高等教育经管类专业系列教材

ISBN 978-7-302-60301-6

I. ①企… II. ①汪… ②金… III. ①企业经营管理－数据处理－高等学校－教材 IV. ①F272.3

中国版本图书馆 CIP 数据核字(2022)第 037384 号

责任编辑：刘金喜
封面设计：周晓亮
版式设计：思创景点
责任校对：成凤进
责任印制：曹婉颖

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编：**100084
社 总 机：010-83470000　　**邮　　购：**010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：北京同文印刷有限责任公司
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张：**18.25　　**字　　数：**456 千字
版　　次：2022 年 4 月第 1 版　　**印　　次：**2022 年 4 月第 1 次印刷
定　　价：68.00 元

产品编号：095984-01

前　　言

随着“大智移云物区”——大数据、人工智能、移动互联网、云计算、物联网和区块链等技术的快速发展，企业越来越多的财务会计工作内容被自动化、智能化所取代。因此，财会人员面临数字化转型，而转型的方向是广大会计人员需思考的问题。Gartner 发布的 2021 年 CFO 调查报告中，CFO 们一致认为应用 RPA、数据分析是重中之重。其中，高级数据分析是财务人员转型的一个重要方向。

【从财务分析到经营分析】

从数据分析内容角度来看，企业财务人员应从财务分析向经营分析过渡。财务分析是财务工作的数据分析，但仅有财务分析是不够的。在企业中，财务分析定位问题，经营分析解决问题，因此，经营分析是针对财务分析发现的问题结合业务所做的更深入的专项分析。过去，由于企业信息化程度不够，企业很难快速收集大量的业务数据并做到及时有效的经营分析，因此，财务分析工作就显得很重要。随着企业信息化程度的加深，各种智能技术的应用，企业内外部信息系统积累了大量的财务业务数据，很多财务数据就是由业务数据驱动生成的。这种情况下，企业在财务分析中发现的问题，可以继续通过经营分析探寻其原因。例如，企业今年管理费用中的差旅费同比大幅度上升，就可以通过与财务系统集成的商旅系统获取的数据进行经营分析，查看业务人员的交通费和住宿费，对数据层层深入挖掘，找到差旅费上升的真正原因。目前，很多大企业都非常重视经营分析，如华为公司就非常推崇经营分析，而不是单纯的财务分析。企业中，财务分析是 CFO 工程，经营分析是 CEO 工程。

从数据分析工具角度来看，有侧重数据采集的 Kettle、Python；有侧重数据仓库的 Hadoop、Teradata 等；有侧重数据挖掘的 SaS、SPSS、R、Python 等；有侧重数据可视化的 Tableau、PowerBI、FineBI、SmartBI 等。这些工具有的需要具备计算机编程知识，有的需要具备统计学知识，因此都不适合财经类人员快速掌握数据分析技能。对于熟悉业务、财务的财会人员来说，从侧重数据可视化的分析工具入手学习掌握数据分析技能是最佳选择。因此，PowerBI、FineBI 等商务智能分析工具是让业务财务人员提升数据分析能力的有效利器。

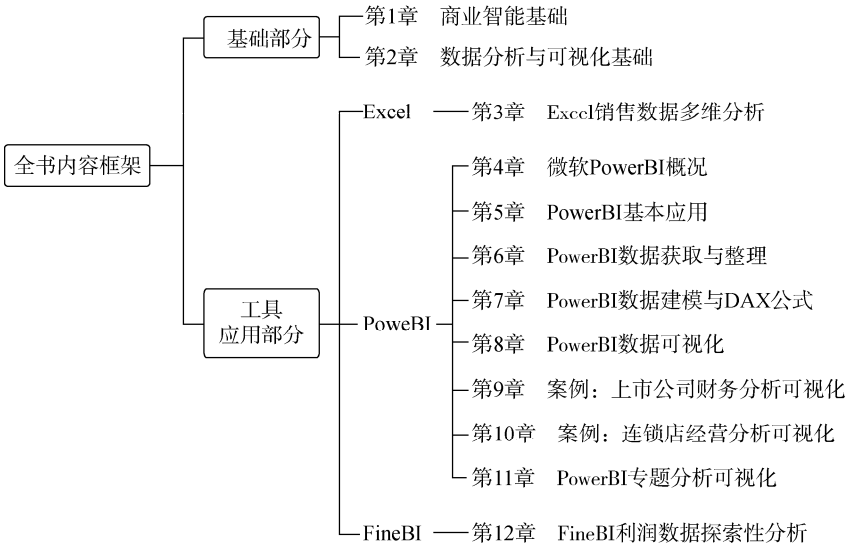
【本书面向对象】

本书作为入门级教材，面向普通高等院校(包括应用型本科、高职)经管类相关专业的学生，以及对数据分析感兴趣的企事业单位人员，旨在使读者了解商业智能及数据分析可视化的基本原理，熟悉数据分析的步骤及思路，掌握 PowerBI、FineBI 等常见自助式 BI 工具的一般应用。

本书对应的高校课程有“智能财务分析可视化”“财会大数据智能分析可视化”“商业智能 BI 在会计中的应用”等。

【本书内容设计】

全书内容框架如下图所示。



全书分为基础部分和工具应用部分：基础部分主要介绍商业智能基础、数据分析与可视化基础等理论知识；工具应用部分介绍 Excel、PowerBI、FineBI 等商业智能数据分析工具，尤其重点介绍 PowerBI 的应用。

1. Excel

Excel是企业各类人员都应掌握的数据分析工具。当很多中小企业数据量不大或不需要做复杂数据分析时，Excel 就能显示其灵活性和适用性。因此，企业各类人员应掌握 Excel 数据分析的一般思路。

2. PowerBI

PowerBI 是微软公司的自助式 BI 工具，其功能全面、强大，尤其是数据处理方面的功能，比其他 BI 工具更多、更有优势。因此，本书重点选择了 PowerBI 这个自助式 BI 工具，详细介绍了其数据获取与整理、数据建模、数据分析与可视化、在线应用等功能。

3. FineBI

FineBI 是帆软公司推出的自助式 BI 工具，在国内市场有很高的占有率。本书通过利润数据探索性分析案例来介绍 FineBI 数据分析的一般思路。

通过以上 3 个工具的安排，希望学员既能掌握基础数据分析工具 Excel 的使用，也能掌握国内外常见自助式 BI 工具的应用。

【本书板块设计】

本书共设计了 5 大板块，方便学员自学，如下表所示。

板块	作用
学习目标	学完本章应了解、熟悉和掌握的知识点
引导案例	每章前都设计一个引导案例。此案例是与本章教学内容相关的模拟案例、实际案例
内容讲解	每章详细内容展开，包括理论介绍、案例数据、操作步骤、小知识、数据分析思维等方面
本章小结	用思维导图对本章知识点进行总结
本章习题	通过 5 种题型(单选题、多选题、判断题、思考题、实训题)的训练，不仅使读者进一步消化、理解本章理论知识，而且能够提升读者的实际操作能力

【本书教学资源】

本书配套了 6 种教学资源，方便教师教学和学生自学，如下表所示。

序号	资源种类	形式	数量	使用对象
1	教学大纲	Word 格式	1 份	教师
2	教案	Word 格式	1 份	教师
3	习题答案	PDF 格式	1 份	教师
4	教学课件	PPT 格式	12 份	教师、学生
5	微课(内容讲解、操作演示)	MP4 格式	若干	教师、学生
6	案例数据	Excel 格式 文本格式 csv 格式 mdb 格式 pbix 格式	若干	教师、学生

上述资源可通过扫描下方二维码下载。微课视频可通过扫描书中二维码观看。为便于开展线上线下混合教学，本书还提供了慕课学习平台，任课老师可发邮件至 476371891@qq.com 获取链接地址。

教学资源下载

【致谢】

本书由汪刚、金春华担任主编，刘畅、舒艳、李洁担任副主编。另外，参加本书编写的人员

还有贾香萍、沈银萱、宁宇、邓宁、吕晓敏、柴明洋、张思涵。

感谢 PowerBI 领域的专家赵文超、宗萌、马世权、高飞、采悟、雷元、武俊敏等老师，感谢他们基于企业实践经验所带来的精彩纷呈的 PowerBI 课程和经验分享。

感谢帆软软件公司、思迈特软件公司提供教学案例及客户案例。

感谢清华大学出版社对本书出版给予的大力支持。

限于作者水平，对于书中的疏忽及错漏之处，诚挚地希望广大读者给予批评指正。

作　者

2021 年 9 月

目　　录

第 1 章　商业智能基础 …… 1
1.1　商业智能概述 …… 2
1.1.1　商业智能的定义 …… 2
1.1.2　商业智能与 DIKW 模型 …… 3
1.1.3　商业智能的系统架构 …… 4
1.1.4　商业智能的价值 …… 5
1.2　商业智能核心技术 …… 6
1.2.1　ETL 技术 …… 6
1.2.2　数据仓库技术 …… 7
1.2.3　联机分析处理技术 …… 8
1.3　自助式商业智能分析工具 …… 10
1.3.1　PowerBI …… 10
1.3.2　Tableau …… 11
1.3.3　FineBI …… 11
1.3.4　SmartBI …… 11
【本章小结】 …… 12
【本章习题】 …… 13

第 2 章　数据分析与可视化基础 …… 15
2.1　数据类型与企业大数据 …… 16
2.1.1　数据类型 …… 16
2.1.2　大数据 …… 18
2.1.3　企业业财大数据 …… 19
2.2　数据分析基础 …… 20
2.2.1　数据分析类型 …… 20
2.2.2　数据分析步骤 …… 20
2.2.3　数据分析主题与框架 …… 22
2.2.4　数据分析模型 …… 24
2.2.5　数据分析指标 …… 30
2.2.6　数据分析方法 …… 33
2.3　数据可视化基础 …… 34
2.3.1　数据可视化原理 …… 34
2.3.2　可视化图表选择 …… 37
【本章小结】 …… 41
【本章习题】 …… 41

第 3 章　Excel 销售数据多维分析 …… 44
3.1　Excel 数据分析基础 …… 44
3.1.1　认识 Excel 数据表 …… 45
3.1.2　相对地址、绝对地址与混合地址 …… 45
3.1.3　Excel 数据分析常用函数 …… 46
3.1.4　Excel 常见错误值及原因 …… 48
3.2　Excel 数据分析基本流程 …… 50
3.2.1　数据获取 …… 50
3.2.2　数据处理 …… 50
3.2.3　数据分析 …… 51
3.2.4　数据可视化 …… 53
3.3　案例：销售数据多维分析可视化 …… 54
3.3.1　打开销售数据表 …… 54
3.3.2　销售数据处理 …… 54
3.3.3　销售数据多维分析 …… 58
3.3.4　销售数据可视化 …… 61
3.3.5　销售数据看板制作 …… 63
3.3.6　销售数据看板变换风格 …… 65

【本章小结】……66
【本章习题】……67
第 4 章　微软 PowerBI 概况……69
4.1　PowerBI 概述……70
4.1.1　传统 BI 与自助式 BI……70
4.1.2　PowerBI 简介……71
4.1.3　PowerBI 特点……72
4.2　PowerBI 应用模式及系列组件……73
4.2.1　PowerBI 应用模式……73
4.2.2　PowerBI 系列组件……74
4.3　PowerBI Desktop 安装与账号注册……75
4.3.1　PowerBI Desktop 安装……75
4.3.2　PowerBI 账号注册……76
4.4　PowerBI Desktop 界面……78
4.4.1　菜单栏……78
4.4.2　视图……78
4.4.3　报表编辑器……80
【本章小结】……82
【本章习题】……82
第 5 章　PowerBI 基本应用……84
5.1　数据获取与整理……86
5.1.1　数据获取……86
5.1.2　数据整理……88
5.2　数据建模……92
5.2.1　建立数据模型……92
5.2.2　新建列……93
5.2.3　新建度量值……94
5.3　数据可视化……95
5.3.1　插入图片、文本框、形状……95
5.3.2　插入卡片图……96
5.3.3　插入环形图……97
5.3.4　插入条形图……97
5.3.5　插入折线图……98
5.3.6　插入气泡图……99
5.3.7　插入切片器……100
5.3.8　图表美化……101
5.4　在线应用……102
5.4.1　在线发布……102
5.4.2　制作仪表板……103
5.4.3　发布到 Web……105
5.4.4　移动应用……105
【本章小结】……106
【本章习题】……107
第 6 章　PowerBI 数据获取与整理……109
6.1　一维表和二维表……110
6.2　数据获取……111
6.2.1　从文件导入……111
6.2.2　从文件夹导入……113
6.2.3　从数据库导入……115
6.2.4　从网站查询导入……116
6.2.5　从其他数据源导入……118
6.2.6　重新设定数据源……119
6.3　数据整理……119
6.3.1　查询编辑器和 M 语言……119
6.3.2　数据的行、列操作和筛选……121
6.3.3　数据类型的转换……126
6.3.4　数据格式的转换……127
6.3.5　数据的拆分、提取和合并……129
6.3.6　数据的转置和反转……133
6.3.7　数据的透视和逆透视……135
6.3.8　分组依据……136
6.3.9　添加列……137
6.3.10　日期和时间的整理……139
6.3.11　数据的基本数学运算……141
6.3.12　数据的组合……143
6.3.13　合并查询……144
【本章小结】……147
【本章习题】……147

第 7 章　PowerBI 数据建模与 DAX 公式 …… 150
7.1　管理关系 …… 151
7.1.1　认识表 …… 151
7.1.2　认识关系及关系模型 …… 151
7.1.3　创建关系 …… 154
7.2　新建列与新建度量值 …… 158
7.2.1　新建列 …… 158
7.2.2　新建度量值 …… 160
7.3　DAX 数据分析表达式 …… 161
7.3.1　认识 DAX 公式 …… 161
7.3.2　认识 CALCULATE 函数 …… 168
7.3.3　认识 DIVIDE 函数 …… 170
【本章小结】 …… 172
【本章习题】 …… 172

第 8 章　PowerBI 数据可视化 …… 175
8.1　常用可视化图表 …… 176
8.1.1　新建仪表图 …… 179
8.1.2　新建百分比仪表图 …… 180
8.1.3　新建表 …… 181
8.1.4　新建矩阵 …… 182
8.2　自定义可视化图表 …… 183
8.2.1　添加自定义可视化对象 …… 183
8.2.2　马表图 …… 184
8.2.3　子弹图 …… 186
8.2.4　文字云 …… 187
8.2.5　桑基图 …… 187
8.3　图表美化 …… 188
8.3.1　切换主题 …… 188
8.3.2　设置图表格式 …… 189
8.4　图表的筛选、钻取和编辑交互 …… 190
8.4.1　图表的筛选 …… 191
8.4.2　图表的钻取 …… 195
8.4.3　图表的编辑交互 …… 196
【本章小结】 …… 198
【本章习题】 …… 198

第 9 章　案例：上市公司财务分析可视化 …… 201
9.1　资产负债表可视化 …… 205
9.1.1　插入切片器 …… 205
9.1.2　插入卡片图 …… 206
9.1.3　插入圆环图 …… 207
9.1.4　插入饼图 …… 208
9.1.5　插入分区图 …… 208
9.1.6　插入树状图 …… 209
9.1.7　插入表 …… 209
9.2　利润表可视化 …… 211
9.2.1　插入卡片图 …… 212
9.2.2　插入圆环图与树状图 …… 212
9.2.3　插入折线图和分区图 …… 212
9.2.4　插入矩阵 …… 212
9.3　现金流量表可视化 …… 213
9.3.1　插入卡片图 …… 214
9.3.2　插入圆环图 …… 214
9.3.3　插入分区图 …… 214
9.3.4　插入折线图 …… 215
9.3.5　插入桑基图 …… 215
9.3.6　插入簇状条形图 …… 215
9.4　偿债能力分析可视化 …… 216
9.4.1　插入卡片图 …… 216
9.4.2　插入折线图 …… 217
9.5　营运能力分析可视化 …… 217
9.5.1　插入卡片图 …… 217
9.5.2　插入折线图 …… 218
9.6　盈利能力分析可视化 …… 218
9.6.1　插入卡片图 …… 218
9.6.2　插入折线图 …… 219
9.7　杜邦分析可视化 …… 219
9.7.1　插入卡片图 …… 219
9.7.2　插入图形图像 …… 220
【本章小结】 …… 220

【本章习题】……221

第 10 章　案例：连锁店经营分析可视化……222

10.1　产品分析可视化……225

10.1.1　插入卡片图……226

10.1.2　插入条形图……226

10.1.3　插入圆环图……226

10.1.4　插入瀑布图……227

10.1.5　插入柱形图……228

10.1.6　插入矩阵……228

10.2　区域分析可视化……229

10.2.1　插入圆环图……230

10.2.2　插入条形图……230

10.2.3　插入柱形图……230

10.2.4　插入水族馆图……230

10.2.5　插入矩阵……230

10.3　趋势分析可视化……231

10.3.1　插入分区图……231

10.3.2　插入折线和柱形图……232

10.3.3　插入动态气泡图……232

10.3.4　插入表……232

10.4　完成度分析可视化……233

10.4.1　插入子弹图……234

10.4.2　插入仪表图……235

10.4.3　插入百分比仪表图……235

10.4.4　插入水平条形图……236

10.4.5　插入矩阵……237

10.5　排名分析可视化……238

10.5.1　插入条形图(排名前 *N* 个)……238

10.5.2　插入表……239

10.5.3　插入文字云……239

【本章小结】……240

【本章习题】……240

第 11 章　PowerBI 专题分析可视化……242

11.1　客户价值 RFM 分析……243

11.1.1　理解数据……243

11.1.2　查看数据模型……244

11.1.3　新建度量值……244

11.1.4　新建列……245

11.1.5　插入切片器……246

11.1.6　插入卡片图……246

11.1.7　插入条形图……247

11.1.8　插入瀑布图……247

11.1.9　插入表……248

11.2　动态帕累托分析……248

11.2.1　理解并连接数据……248

11.2.2　新建度量值……249

11.2.3　插入切片器……250

11.2.4　插入表……250

11.2.5　插入圆环图……250

11.2.6　折线和堆积柱形图……251

11.3　盈亏平衡动态分析……252

11.3.1　理解数据……252

11.3.2　获取数据……252

11.3.3　创建数据表……252

11.3.4　创建度量值……253

11.3.5　插入切片器……254

11.3.6　插入分区图……254

11.4　杜邦动态分析……255

11.4.1　插入卡片图……255

11.4.2　设置变化参数……256

11.4.3　修改度量值……257

【本章小结】……258

【本章习题】……258

第 12 章　FineBI 利润数据探索性分析……259

12.1　FineBI 简介……260

12.1.1　FineBI 概述……260

12.1.2　FineBI 的安装与启动……261

12.1.3　FineBI 的界面……262

12.2　FineBI 数据分析基本流程……263

12.3　案例：FineBI 利润数据探索型分析可视化 ……263
12.3.1　案例背景 ……263
12.3.2　数据准备 ……264
12.3.3　数据加工 ……265
12.3.4　仪表板制作及可视化分析 …… 267
【本章小结】 …… 277
【本章习题】 …… 277

第 1 章

商业智能基础

学习目标

- 了解商业智能的定义、价值;
- 了解自助式商业智能分析工具;
- 熟悉DIKW模型和商业智能的系统架构;
- 掌握商业智能核心技术ETL、数据仓库、联机数据处理OLAP。

引导案例

啤酒和尿布的故事

这是发生在美国沃尔玛连锁超市的真实案例，并一直为商家所津津乐道。沃尔玛拥有世界上最大的数据仓库系统，为了能够准确了解顾客在其门店的购买习惯，沃尔玛对其顾客的购物行为进行购物篮分析。沃尔玛数据仓库里集中了其各门店的详细原始交易数据，在这些原始交易数据的基础上，沃尔玛利用数据挖掘方法对这些数据进行分析和挖掘。分析结果意外地发现：与尿布一起购买最多的商品竟是啤酒！经过大量实际调查和分析，揭示了一个隐藏在“尿布与啤酒”背后的美国人的一种行为模式：在美国，一些年轻的父亲下班后经常要到超市去买婴儿尿布，而他们中有30%～40%的人同时也会为自己买一些啤酒。产生这一现象的原因是：美国的太太们常叮嘱她们的丈夫下班后为小孩买尿布，而丈夫们在买尿布后又随手带回了他们喜欢的啤酒。

这个著名的“啤酒和尿布的故事”，是商业智能领域的经典案例。沃尔玛正是根据这一发现，及时调整营销策略，在尿布旁边摆放啤酒，从而大大增加了两种商品的销售收入。

1.1 商业智能概述

现代管理之父彼得·德鲁克(Peter F. Drucker)曾经说，知识已经成为关键的经济资源和竞争优势的主要来源。如何将数据转化为知识？这就需要商业智能(business intelligence，BI，也叫商务智能)发挥作用。商业智能是将不可用数据转变为可行的见解的过程。商业智能的出现，被认为是继 ERP 之后企业信息化的又一个热潮。商业智能把各种数据及时地转换为支持决策的信息和知识，帮助企业管理者了解客户的需求、消费习惯，预测市场的变化趋势及行业的整体发展方向，进行有效的决策，从而在竞争中占据有利地位。

1.1.1 商业智能的定义

商业智能从 20 世纪 90 年代开始，已经在众多企业中引起了广泛关注，成为业界关注的热点。商业智能的概念最早由美国著名咨询公司 Gartner 的分析师霍华德·德雷斯纳(Howard Dresner)于 1996 年提出，他认为商业智能描述了一系列的概念和方法，应用基于数据的分析系统辅助商业决策的制定。商业智能技术为企业提供了迅速收集、分析数据的技术和方法，把这些数据转化为有用的信息，提高了企业决策的质量。

企业界对商业智能存在着不同的理解，主要如下。

Microsoft 认为：商业智能是任何尝试获取、分析企业数据以便更清楚地了解市场和顾客，改进企业流程，更有效地参与竞争的过程。

Oracle 认为：商业智能是一种商务战略，能够持续不断地对企业经营理念、组织结构和业务流程进行重组，实现以顾客为中心的自动化管理。

SAP 认为：商业智能是收集、存储、分析和访问数据以帮助企业更好决策的技术。

IBM 认为：商业智能是一系列技术支持的简化信息收集、分析的策略集合。

IDC 认为：商业智能是终端用户查询和报告工具、在线分析处理工具、数据挖掘软件、数据集市、数据仓库产品等软件工具的集合。

帆软认为：商业智能是利用数据仓库、数据可视化与分析技术，将指定的数据转化为信息和知识的解决方案，其价值体现为满足企业不同人群对数据查询、分析和探索的需求，使企业实现对业务的监测和洞察，从而支撑企业管理决策，提升企业管理水平，提高企业业务运营效率，改进优化企业业务。

综上可以看出，商业智能是基于先进信息技术与创新管理理念，集成了企业内外各种类型数据，对其加工处理并从中提取能够创造商业价值的信息和知识，面向企业战略并服务于管理层、业务层，指导企业经营决策，提升企业竞争力。商业智能不仅是指一套技术，更是一套完整的解决方案，商业智能的本质就是数据智能。这里所说的数据包括：来自企业业务系统的订单、交易、库存等数据；来自企业财务系统的凭证、账簿、报表等数据；来自企业外部的客户和供商数据、竞争对手数据、企业所处其他外部环境中的各种数据。

1.1.2　商业智能与DIKW模型

DIKW 金字塔模型是知识管理领域中非常重要的模型，它将数据、信息、知识、智慧纳入一种金字塔形的层次体系，每一层比下一层都赋予了新的特质。DIKW金字塔模型如图 1-1 所示。

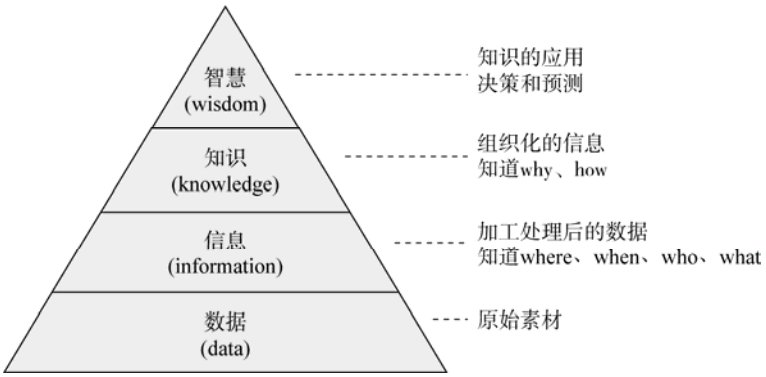


图1-1　DIKW金字塔模型

从该模型可以看出：数据可以是数字、文字、图像、符号等，它直接来自于事实，可以通过原始的观察或度量来获得。数据可以是定量的，也可以是定性的。信息是已经被处理并具有逻辑关系的数据，是对数据的解释，这种信息对其接收者具有意义。这些信息可以回答一些简单的问题，如谁？什么？哪里？什么时候？因此，信息也可以看成是被理解了的消息。知识是从相关信息中过滤、提炼及加工而得到的有用资料，它体现了信息的本质、原则和经验。此外，知识基于推理和分析，还可能产生新的知识。知识可以回答“如何？”的问题，可以帮助我们建模和仿真。智慧是人类所表现出来的一种独有的能力，主要表现为收集、加工、应用、传播知识的能力，以及对事物发展的前瞻性看法。在知识的基础上，通过经验、阅历、见识的累积，而形成的对事物的深刻认识、远见，体现为一种卓越的判断力。通过 DIKW 模型分析，可以看到数据、信息、知识与智慧之间既有联系，又有区别。

我们从本章引导案例“啤酒和尿布的故事”中来看一看从数据到智慧的转化过程。沃尔玛超市 POS 机中某种商品的销售记录，包括客户ID、商品名称、销售数量、销售金额等，这就是数据。沃尔玛超市将所有销售记录加工处理后形成了销售信息，如不同商品的销售金额信息、不同店铺的销售金额信息等。沃尔玛超市通过进一步分析和挖掘数据，发现啤酒和尿布的销量成正相关关系，这就是知识。根据啤酒和尿布的销量成正相关关系这个知识，沃尔玛公司及时调整销售策略，将两种商品摆在相近的位置进行销售，从而大大提高了两种商品的销量，这就是智慧。

DIKW 金字塔模型可以展现从数据到智慧的转化过程，那商业智能和此模型有什么关系呢？在大数据面前，通过手工处理来实现从数据到智慧的转化过程是很难做到的，或者说时间会很长，人们在做出决策采取行动时，往往错过了时效。商业智能系统的出现，可以帮助从数据到知识的有效转化，提升转化准确率，可以加快从数据到智慧的转化时间，使人们及时做出决策并采取行动。

1.1.3 商业智能的系统架构

商业智能需要将企业积累的大量数据处理成信息，再转化为知识，最后通过可视化方式将信息和知识展现给企业相关人员，便于企业进行商务决策。商业智能架构体系分为数据获取、数据管理、数据分析、数据展示 4 层，如图 1-2 所示。

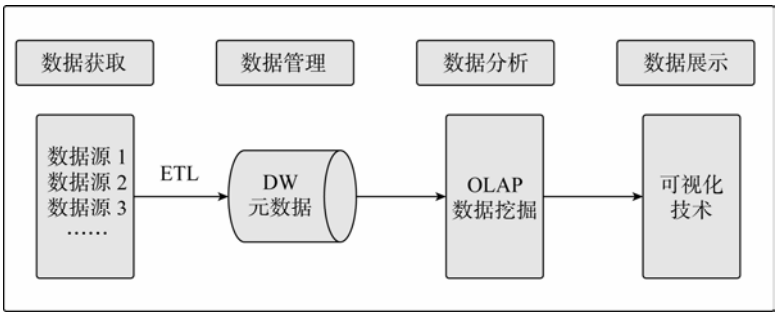


图1-2　商业智能架构体系

1. 数据获取层

商业智能是基于数据的智能，商业智能的基础是数据。获取什么样的数据及如何获取数据是商业智能系统数据获取层的主要工作内容。

从来源上看，数据通常包括企业内部的数据和企业外部的数据。企业内部的数据又包括来自财务系统、采购系统、销售系统、生产系统等 ERP 系统的数据、办公自动化 OA 的数据；企业外部的数据包括政策数据、市场数据、行业数据、竞争对手的数据、各类统计数据等。从存储结构上看，数据分为结构化数据和非结构化数据。结构化数据主要是以二位表格形式存储的数据，非结构化数据主要包括所有格式的办公文档、文本、图片、音频和视频文件等。从存储形式上看，数据分为.xls 文件、.txt 文件、.csv 文件、各类数据库文件等。目前的 BI 系统都能直接读取并连接各种类型的结构化数据。

BI 系统获取数据后，还要经过数据处理，即 ETL(extract-transform-load)操作。

2. 数据管理层

数据管理层主要通过数据仓库(data warehouse，DW)和元数据(meta data)管理方式实现对数据的管理。

经过 ETL 处理的数据被加载到数据仓库中。数据仓库是面向主题的，数据仓库中的数据按照一定的主题域进行组织。数据仓库中的数据是在对原有分散的数据库数据抽取、清理的基础上经过系统加工、汇总和整理得到的，需消除源数据中的不一致性，以保证数据仓库内的信息是关于整个企业的、有效的、可用的信息。数据仓库中的数据通常是历史数据，通过这些数据，可以对企业的过去做分析，对未来做预测。数据仓库所涉及的数据操作主要是数据查询，一旦某个数据进入数据仓库以后，一般情况下将被长期保留，也就是数据仓库中一般有大量的查询操作，但修改和删除操作很少，通常只需要定期加载、刷新。数据仓库存储模型主要包括概念模型、逻辑模型和物理模型。

为了便于管理数据仓库中的数据引入了元数据的概念。元数据是关于数据仓库的数据，指在数据仓库建设过程中所产生的有关数据源定义、目标定义、转换规则等相关的关键数据。同时元数据还包含关于数据含义的商业信息，所有这些信息都应当妥善保存，并很好地管理，为数据仓库的发展和使用提供方便。

3. 数据分析层

数据分析层主要包括库联机分析处理(online analytical processing，OLAP)和数据挖掘(data mining，DM)两种分析。它们都是数据库或数据仓库的分析工具，两者相辅相成，是决策分析不可缺少的工具。

联机分析处理(OLAP)是以海量数据为基础的复杂分析技术。它支持各级管理决策人员从不同的角度，快速灵活地对数据仓库中的数据进行复杂查询和多维分析处理，并且能以直观易懂的形式将查询和分析结果展现给决策人员。OLAP 使用的逻辑数据模型为多维数据模型。数据挖掘是从海量数据中，提取隐含在其中的、人们事先不知道但又可能有用的信息和知识的过程。数据挖掘的数据有多种来源，包括数据仓库、数据库或其他数据源。

可以看出，OLAP 是验证型的分析工具，而数据挖掘是预测型的分析工具。

4. 数据展示层

数据展示层主要是通过可视化技术将分析内容以各种图表的方式展示出来，供企业决策人员、管理人员、分析人员、业务人员等相关人员进行洞察和决策。可视化技术是以计算机图形学、图像处理技术为基础，将数据转换为图形或图像形式，显示到屏幕上，并进行交互处理的理论方法和技术。它涉及计算机视觉图像处理、计算机辅助设计、计算机图形学等多个领域，并逐渐成为一种研究数据表达，综合处理决策分析的问题的综合技术。数据分析的内容包括利润分析、收入分析、成本分析、资产分析、运营分析、投融资分析等。企业数据分析所需要展示的指标和内容与企业战略、经营管理需要有密切关系。

1.1.4　商业智能的价值

商业智能可以帮助管理者减少收集、处理信息的时间，把更多的精力用于决策上。商业智能的价值具体体现在以下几个方面。

1. 增强业务洞察能力

商业智能可减少经营者收集数据、获取信息所花费的时间，加速决策过程，使正确的信息在正确的时间在信息系统中流向各类相关人员。决策者通过监控关键绩效指标(key performance indicator，KPI)，掌控业务执行的状况，以便及时调整策略。例如，管理者通过 KPI 监控销售人员最新销售信息、任务额信息、任务完成度信息，可随时掌握企业营业收入完成情况。

2. 优化企业营销策略

企业通过构建商业智能分析模型，挖掘消费者行为，制定适当的营销策略。例如，“啤酒和尿布的故事”，就是著名的零售企业沃尔玛公司通过商业智能，发现了尿布销售额和啤酒销

售额具有相关性的事实，随后管理层做出决策，将尿布和啤酒两种商品摆放在一起销售，从而提高了企业的销售业绩。

3. 提高市场响应能力

企业借助商业智能的大数据整合能力，将行业信息、政策法规等信息融入商业智能系统，通过适当的模型以预测市场变化，精简流程，确定需要改进的环节，以适应外部环境的变动。

4. 加强风险管理能力

企业可通过商业智能风险预警模型，发现存在的潜在风险点，如经营风险、财务风险、纳税风险等，当出现这些风险预警时，可随时调整企业的经营策略来应对、规避、降低企业的各类风险。例如，银行的贷款业务，应用数据挖掘技术可以对顾客进行信用分析，发现其中的欺诈行为特征，作为有效的预警机制，为银行减少损失。

5. 改善顾客关系管理

现在很多企业正在逐渐由以产品为中心转化为“以顾客为中心”，而应用商业智能中的在线分析处理和数据挖掘等技术，即可通过对顾客的交易记录等相关资料的处理与挖掘，对顾客行为进行分类，然后针对不同类型的顾客制定相应的服务策略，我们把这类应用叫作“顾客智能”。电信企业利用分析模型对顾客行为、信用度等进行评估，对不同类型的客户提供有针对性的服务，从而提高顾客的满意度和忠诚度。

1.2 商业智能核心技术

商业智能(BI)作为一种数据解决方案，能够通过多种技术的综合运用，从大量的数据中筛选出有效的信息，供决策者使用。BI 以大数据为基础，经过 ETL 处理后，将数据按一定方式存储在数据仓库(DW)中，再经过联机分析处理 OLAP 后，即可以发掘出数据的价值。

1.2.1 ETL技术

ETL 是将业务系统的数据经过抽取、转换之后加载到数据仓库的过程，目的是将企业中分散、零乱、标准不统一的数据整合到一起，为企业的决策提供分析依据。ETL 是 BI 项目中的一个重要环节。ETL 的过程如图 1-3 所示。

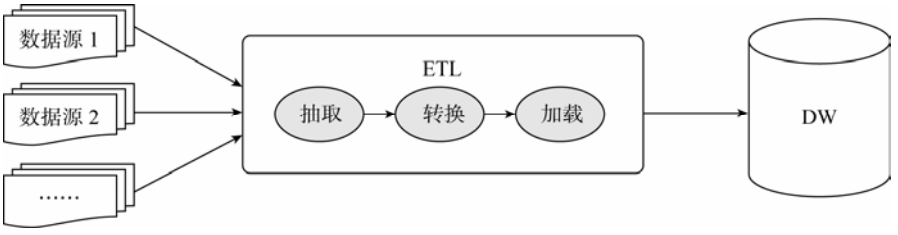


图1-3 ETL的过程

1. 数据抽取(E-extract)

数据仓库中的数据是面向主题的，在将源数据导入数据仓库之前，需要识别出与分析决策相关的数据，抽取过程通常考虑以下几个方面：①确定源数据及其含义；②进行数据抽取，确定访问哪些数据库、文件和表；③确定抽取频率；④确定抽取数据保存的位置；⑤无法抽取时的异常处理。

2. 数据转换(T-transform)

当抽取到数据仓库中的数据时，会有缺失或重复等情况。数据转换就是将抽取的数据变成数据分析所需要的、完整的、准确的目标数据，主要包括数据清洗、格式转换、汇总计算等。数据清洗将从数据的准确性、完整性、一致性、唯一性、适时性、有效性几个方面来处理数据的丢失值、越界值、不一致代码、重复数据等问题。

1) 缺失值的处理方法

缺失值即是不完整数据。大多数情况下，缺失值必须手工补充填入(即手工清理)并与数据提供方进行确认。当无法得到数据提供方确认时，某些缺失值可以从本数据源或其他数据源推导出来，这就可以用平均值、最大值、最小值或更为复杂的概率估计代替缺失的值。如果缺失值很少且不影响数据分析，也可将缺失值记录直接删除。

2) 错误值的处理方法

错误值的产生大多是由于业务系统不够健全，在接收并输入数据后没有进行判断直接写入后台数据库造成的，如数值数据输成全角数字、字符串包含若干空格、日期格式不正确、日期越界等。当产生错误值时，可以人工修改或用统计分析的方法识别可能的错误值或异常值，从而将其加以修正。

3) 重复值的处理方法

数据库中属性值相同的记录被认为是重复记录，通过判断记录间的属性值是否相等来检测记录是否相等，重复的记录只保留一条即可，将多余的记录删除。

4) 不一致性数据的处理方法

从多数据源集成的数据可能有语义冲突，例如，性别字段中的“男性”，有的表中表示为 1，有的表中表示为“男”。当出现不一致性数据时，可定义完整性约束用于检测不一致性，也可通过分析数据发现联系，从而使数据保持一致。

3. 数据加载(L-load)

数据加载是将清洗和转换后的、符合数据分析要求的数据加载到数据仓库中，加载方式包括直接追加和全面覆盖。

1.2.2　数据仓库技术

数据仓库(DW)是一个面向主题的、集成的、相对稳定的、反映历史变化的数据集合，用于支持管理决策。20 世纪 80 年代中期由威廉・汉・荫蒙(William H.Inmon)首次提出。

数据仓库模型主要包括概念模型、逻辑模型和物理模型。概念模型通过主题来表达，用维和度量表示。逻辑模型也叫关系模型，用于确定关系模式的定义，如数据分割策略。物理模型是逻辑模型在数据仓库中的实现，如数据索引策略、数据的存储策略及存储优化分配等。下面我们重点介绍逻辑模型。

逻辑模型一般有两种，即星型模型和雪花模型。

1. 星型模型

星型模型由事实表和维度表组成，事实表可连接多种维度表，维度表只有一层。维度表和事实表必须有能够关联的字段。例如，某公司采购业务的数据星型模型如图 1-4 所示。其中：采购订单表是事实表，包含企业实际的全部采购订单数据，通常记录数很多；而日期表、商品表、物流表、供应商表等都是维度表，其中的数据作为分析的维度，而且数据量很少。

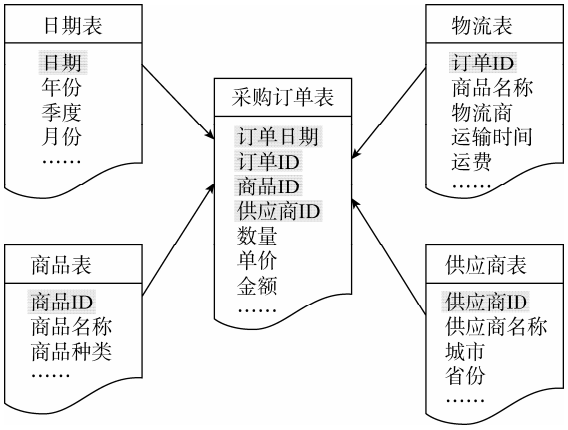


图1-4　某公司采购业务的数据星型模型

2. 雪花模型

雪花模型是星型模型的扩展，在事实表的外部有多层维度表。例如，将供应商表中的城市和省份字段拆分出来放入地区表中，再将地区表作为供应商表的维度表，这样做，可以减少数据冗余。

在数据联机分析 OLAP 中，使用星型模型的情况较多。

1.2.3　联机分析处理技术

联机分析处理(OLAP)技术是使用最广的数据分析技术，其通过对数据仓库的多维分析，可以快速实现洞察并发现问题。

OLAP 中最主要的操作结构是数据立方体，它是一种支持快速数据分析的多维数据结构(实际的或虚拟的)，能够进行多维度高效率操作和数据分析，如图 1-5 所示。使用 OLAP，分析人员通过改变数据位置和定义来分析，可以在整个数据库中进行导航，并能提取数据的一个特定子集。常见的 OLAP 操作包括切片、切块、钻取、上卷和旋转等。

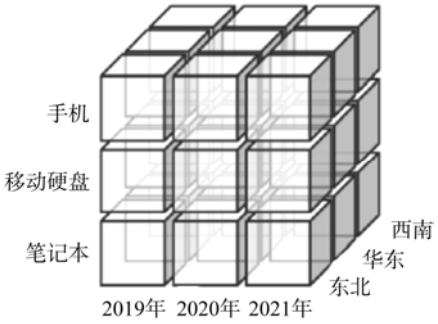


图1-5　数据立方体

1. 切片(slice)

切片操作是选择特定的维度值进行分析。例如，只选择东北地区的销售数据，或者只查看2021年度的销售数据，如图 1-6 所示。

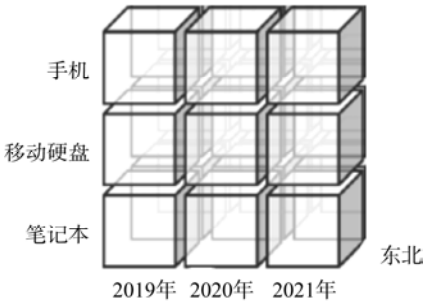


图1-6　切片

2. 切块(dice)

切块操作是选择维度中特定区间的数据或某批特定值进行分析。例如，选择 2020—2021年的销售数据，如图 1-7 所示。

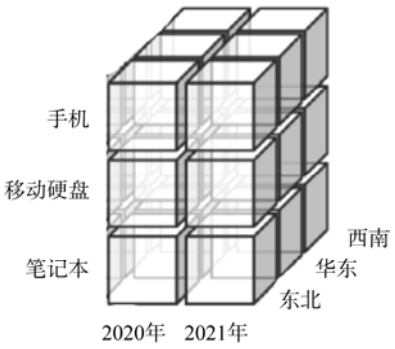


图1-7　切块

3. 钻取(drill-down)

钻取操作是在维度的不同层次间的变化，如从上层维度降到下层维度，或者将汇总数据拆分为更细节的数据。例如，对手机维度向下钻取，得到的数据立方体如图 1-8 所示。

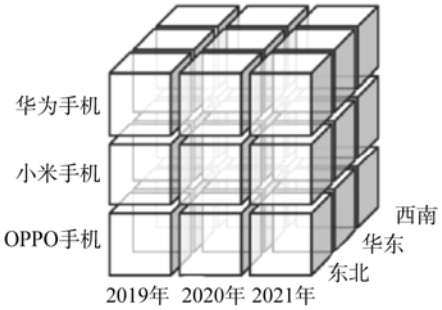


图1-8　钻取

4. 上卷(roll-up)

上卷操作是钻取的逆操作，即从低层维度向高层维度聚合。例如，将华为手机、小米手机和 OPPO 手机向上汇总。

5. 旋转(pivot)

旋转即维度位置的互换，就像是二维表的行列转换。例如，将年份维度和地区维度互换，如图 1-9 所示。

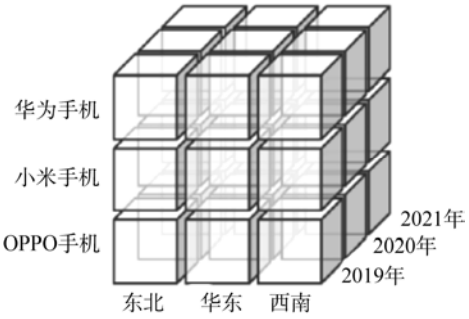


图1-9　旋转

1.3　自助式商业智能分析工具

自助式 BI 不再面向 IT 部门的技术人员，而是面向不具备 IT 背景的业务人员。与传统 BI 相比，自助式 BI 更灵活且更易于使用。在当下，越来越多的业务人员，基于自身对业务的理解，使用自助式 BI 系统，能够快速实现数据分析与洞察，提高了决策的相应速度和时效性。下面介绍几款自助式 BI 分析工具。

1.3.1　PowerBI

PowerBI 是微软官方推出的可视化数据探索和交互式报告工具。PowerBI 能让静态数据转化为动态报表，是一个让非专业数据分析人员也能做到有效整合企业数据，并快速准确地提供商业智能分析的数据可视化工具。

2021 年 2 月 18 日，国际著名咨询机构 Gartner 公司发布的《商业智能和分析平台魔力象限》(*Magic Quadrant for BI and Analytics Platforms*)年度报告显示，微软连续第 14 年入选，并超越一切对手，再次成为最具领导力和超前愿景的 BI 公司。

PowerBI 应用包含 Windows 桌面应用程序(称为 PowerBI desktop)、联机 SaaS(软件即服务)服务(称为 PowerBI online-service)及移动 PowerBI 应用(称为 PowerBI mobile)。其中桌面应用程序为免费版。

1.3.2　Tableau

Tableau是一个可视化分析平台，它改变了使用数据解决问题的方式，使个人和组织能够充分利用自己的数据。Tableau分析平台作为现代商业智能市场的领先产品使人们能够更加轻松地探索和管理数据，更快地发现和共享可以改变企业和世界的见解。

Tableau 成立于 2003 年，是斯坦福大学一个计算机科学项目的成果，该项目旨在改善分析流程并让人们能够通过可视化更轻松地使用数据。自成立以来，Tableau一直快速、不断地进行研发投资，开发各种解决方案来帮助所有需要使用数据的人更快地找到答案，以发现意想不到的见解。Tableau 在 2019 年被 Salesforce 收购，但使命不变：帮助人们查看并理解自己的数据。

Tableau 家族产品包括 Tableau Desktop、Tableau Server、Tableau Online、Tableau Public 和 Tableau Reader。

Tableau以其简单易用、极速高效、美观交互视图、轻松实现数据融合等优势，帮助人们使用数据推动变革。

1.3.3　FineBI

FineBI 是帆软软件有限公司推出的一款商业智能(business intelligence)产品，它可以通过最终业务用户自主分析企业已有的信息化数据，帮助企业发现并解决存在的问题，协助企业及时调整策略以做出更好的决策，增强企业的可持续竞争性。FineBI 定位于自助大数据分析的 BI 工具，能够帮助企业的业务人员和数据分析师，开展以问题导向的探索式分析。

FineBI 的产品优势是：让业务人员/数据分析师自主制作仪表板，进行探索分析。数据取于业务，用于业务，让需要分析数据的人，可以自己处理分析数据。

FineBI的系统构架包括以下四部分：①数据处理，即数据处理服务，用来对源数据进行抽取、转换、加载，为分析服务生成数据仓库 FineCube；②即时分析，可以选择数据快速创建表格或图表以使数据可视化、添加过滤条件筛选数据，即时排序，使数据分析更快捷；③多维度分析，OLAP 分析实现，提供各种分析挖掘功能和预警功能，如任意维度切换、添加、多层钻取、排序、自定义分组、智能关联等；④Dashboard，提供各种样式的表格和多种图表服务，配合各种业务需求展现数据。

1.3.4　SmartBI

SmartBI 是思迈特软件公司旗下的产品。思迈特软件成立于 2011 年，致力于为企业客户提

供一站式商业智能解决方案。SmartBI 是企业级商业智能 BI 和大数据分析品牌，满足用户的企业级报表、数据可视化分析、自助分析平台、数据挖掘建模、AI 智能分析等大数据分析需求；致力于打造产品销售、产品整合、产品应用的生态系统；与上下游厂商、专业实施伙伴和销售渠道伙伴共同为最终用户服务；通过 SmartBI 应用商店(BI+行业应用)为客户提供场景化、行业化数据分析应用。SmartBI 软件在国内 BI 领域处于领先地位，产品广泛应用于金融、政府、制造、零售、地产等众多行业。

SmartBI 产品系列主要包括以下四大平台。

1) 大数据分析平台

大数据分析平台对接各种业务数据库、数据仓库和大数据分析平台，进行加工处理、分析挖掘和可视化展现；满足所有用户的各种数据分析应用需求，如大数据分析、可视化分析、探索式分析、复杂报表、应用分享等。

2) 数据化运营平台

数据化运营平台围绕业务人员提供企业级数据分析工具和服务满足不同类型的业务用户，在 Excel 或浏览器中都可实现全自助的数据提取、数据处理、数据分析和数据共享，具有很强的适用性。

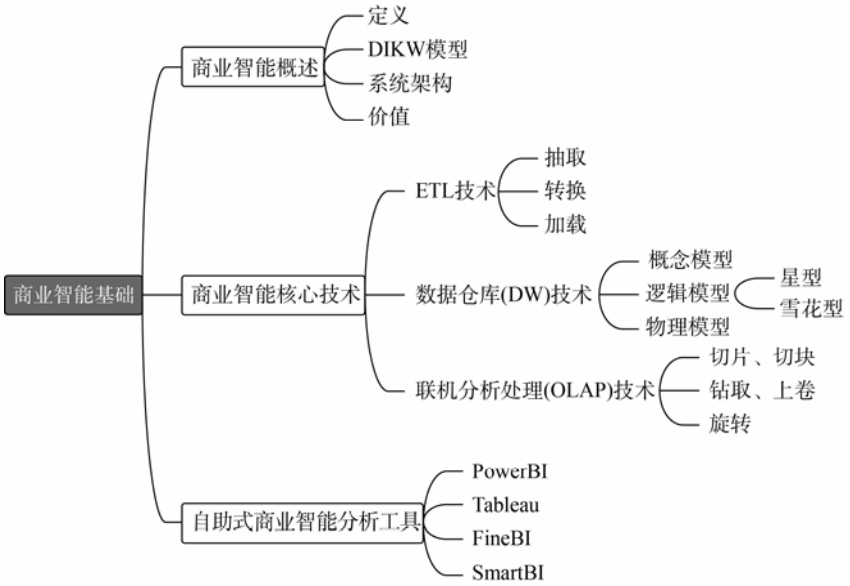
3) 大数据挖掘平台

大数据挖掘平台通过深度数据建模，为企业提供预测能力，支持文本分析、五大类算法和数据预处理，并为用户提供一站式的流程式建模、拖曳式操作和可视化配置体验。

4) SaaS 分析云平台

SaaS(software as a service，软件即服务)分析云平台是全新一代云端数据分析平台，提供自助、快速搭建数据分析应用，可分享深刻见解，提升团队数据智慧。

【本章小结】



【本章习题】

一、单选题

1. 联机分析处理的英文简称是(　　)。
   A. OLAP　　B. OLDP　　C. DM　　D. DW
2. 商业智能的本质是(　　)。
   A. 技术智能　　B. 数据智能
   C. 人才智能　　D. 机器智能
3. 某销售数据分析中，从“地区”维度入手，查看不同省份的销售额，此操作属于 OLAP 的(　　)操作。
   A. 切片　　B. 钻取　　C. 上卷　　D. 切块
4. ETL 中的 T 代表的含义是(　　)。
   A. 数据抽取　　B. 数据转换
   C. 数据加载　　D. 数据清洗
5. 下列说法错误的是(　　)。
   A. 星型模型和雪花模型都属于物理模型
   B. 雪花模型有多层维度表
   C. 数据仓库 DW 是一个面向主题的数据集合
   D. 星型模型有一层维度表

二、多选题

1. 一个完整的 BI 系统的构架包括(　　)。
   A. 数据获取层　　B. 数据管理层
   C. 数据分析层　　D. 数据展示层
2. 商业智能的价值体现在(　　)。
   A. 增强业务洞察能力　　B. 优化企业营销策略
   C. 加强风险管理能力　　D. 改善顾客关系管理
3. 下列属于国产自助式 BI 工具的有(　　)。
   A. FineBI　　B. PowerBI
   C. SmartBI　　D. Tableau
4. DIKW 模型是将(　　)纳入金字塔形的层次体系。
   A. 数据　　B. 信息
   C. 知识　　D. 智慧
5. 数据仓库模型主要包括(　　)。
   A. 物理模型　　B. 概念模型
   C. 层次模型　　D. 逻辑模型

三、判断题

1. 数据仓库的英文简称是 DM。 (　　)
2. DIKW 金字塔模型可以展现从智慧到知识的转化过程。 (　　)
3. 数据仓库中的逻辑模型包括星型和雪花型。 (　　)
4. 星型模型由事实表和维度表组成，事实表可连接多种维度表，维度表可以有多层。 (　　)
5. 切块(dice)是选择维度中特定区间的数据或某批特定值进行分析。 (　　)

四、思考题

1. 请说一说商业智能的系统构架。
2. 商业智能的价值包括哪几个方面？
3. ETL 技术的内容有哪些？
4. 请举例说明星型和雪花模型有什么区别？
5. 联机分析处理技术常见操作包括哪些？

五、实训题

请从 PowerBI、Tableau、FineBI、SmartBI 相应网站上查找一个商业智能应用的典型案例，仔细研究该案例，并将研究成果制作成 PPT 与大家分享。

建议：可以从以下几个方面制作 PPT。

- 项目背景
- 项目目标
- 技术构架
- 建设方案
- 项目价值

以上框架只做参考，可根据研究案例的特点自行设计框架结构。