# 酒店行业BI品牌推荐

来源：网络 作者：心旷神怡 更新时间：2025-07-20

*第一篇：酒店行业BI品牌推荐全球经济一体化进程使得酒店客源日益丰富多样化，加上越来越多的新品牌加入、新酒店开业，使得整个行业面临着日趋激烈的竞争环境和不断攀升的客户期望，迫使经营者不断寻求扩大销售机会、改进服务质量、提升客户满意度以增强酒...*

**第一篇：酒店行业BI品牌推荐**

全球经济一体化进程使得酒店客源日益丰富多样化，加上越来越多的新品牌加入、新酒店开业，使得整个行业面临着日趋激烈的竞争环境和不断攀升的客户期望，迫使经营者不断寻求扩大销售机会、改进服务质量、提升客户满意度以增强酒店的核心竞争力。酒店是一个以服务为本的行业，依靠客人对各项服务的满意度来提升酒店的入住率和经营效益。随着这些竞争的加剧，酒店的每一项服务都会关系着客人对于酒店的评价。

从酒店行业的企业自身来讲，由于酒店客户的流动性较大，所以，需要不断的提管理能力提高企业的效率。近几年，酒店行业的大力发展，产生了大量的数据，但这些数据在酒店企业并没有得到充分的发展与应用，酒店行业的老板需要相应的工具或者手段来把企业的数据转换成最大化的价值。针对于酒店行业企业的需求，商业智能(BI)技术，开始逐渐进入到企业视野。

针对于酒店行业的特殊性以及企业对于商业智能的需求信息，IT168记者推荐了行业内主流的商业智能服务商，帮助酒店企业的老板在选型商业智能提供参考。

一、微软Office Excel

信誉等级：☆☆☆☆☆☆

主流产品: Office Excel

成功案例: 几乎所有有需求的办公人员。

产品评价:简单易用、价格低廉、适用性广，但相对复杂的报表则操作比较困难。行业属性：不限行业属性

企业规模：不受限制

产品参数：Office Excel

注意事项：

处理简单的报表，应用excel比较适用，适合于中小企业创业初期，但随着业务的加大以及数据的复杂性增加，建议用更加专业的报表工具。

产品费用: 费用比较适中

二、奥威智动Power-BI

信誉等级：☆☆☆☆☆

主流产品：Power-BI

产品评价：奥威智动专注为中国企业提供商业智能方案，Power-BI是其近8年的研究成果。定位于传统BI，包括了各项主流BI技术，功能完整强大。平台开发简单，操作灵活，符合大家使用EXCEL的习惯。目前市场范围主要是国内中小企业，性价比较高。行业属性：通用型BI工具

企业规模：中小企业

产品参数：Power-BI

产品费用：费用根据实际的情况衡量，一般在几十万范围内，不限用户数，一次性消费购买。

三、用友华表

信誉等级：☆☆☆☆

主流产品：用友BQ

产品评价：用友收购尚南后，成为用友华表的主打产品。凭借用友在国内的知名度，BQ知名度迅速上升。BQ直接内嵌EXCEL和WORD，设计和展示中国式报表非常有竞争力。但无独立数据仓库，OLAP多维分析功能不够实用。

行业属性：通用型

企业规模：中型或者大型企业

产品参数：用友BQ

产品费用：BQ报价及实施费用较高，并且限制站点。中小企业比较难消费。

四、北京润乾软件技术公司

信誉等级：☆☆☆☆

主流产品: 润乾报表

产品评价：国内知名报表工具。

行业属性：

企业规模：适合中小型企业

产品参数：

产品费用：报表工具，费用根据实际情况评估，可低至几万。

五、文雅科信息技术(上海)有限公司

信誉等级：☆☆

主流产品：商务智能软件Dr.Sum

成功案例：主要的案例集中在日本，国内案例较少。

产品评价：产品在日本应用不错，有许多的案例，但目前在国内应用较少，国内用户对于其产品并不是很熟悉。

行业属性：企业规模：适合中小型企业应用

产品参数：

1)、商务智能软件Dr.Sum

2)、报表工具软件SVF

注意事项：

目前国内案例较少

产品费用：暂无

六、arcplan公司

信誉等级：☆☆☆☆

主流产品: arcplan企业版

产品评价:

行业属性：国内烟草行业应用居多。而在国外涉及通信、金融等领域。

企业规模：适合中小型企业

产品参数：

1)、arcplan企业版

2)、CFO Cockpit

CFO Cockpit是arcplan扩展思维解决财务报表和分析，它的两个版本第一标准(Premier)和HFM给繁忙的、需要快速便捷地完成、实时绩效衡量标准的行政人员提供了许多信息。

3)、arcplan Edge

arcplan Edge是一个强大的，为那些渴望简化流程，缩短预算编制周期的企业提供预算编制，规划和预测(BP&F)的解决方案

4)、arcplan Analytic Services

通过arcplan的Analytic Services(分析服务)，你可以在任何地方实现以业务流程为驱动的分析。

注意事项：烟草行业的客户居多，主要要前端的报表展现。

产品费用：暂无

七、TIBCO Spotfire

信誉等级：☆☆☆☆

主流产品: TIBCO Spotfire Analytics

成功案例: 暂无

产品评价:

与传统的商业智能分析软件相比，Spotfire软件可以形象化地表现信息处理的结果，并能对未来的发展趋势进行预测性分析。该软件采用内存处理模式和独特的形象化互动性数据分析手段，可以直观反映信息处理的结果，便于用户查看和理解。

行业属性： 制造业

企业规模：暂无

产品参数：暂无

注意事项：产品还是比较适用于中小企业，操作简单比较容易易学，但因国外产品相对来讲费用会比国内的产品要略高一些。

产品费用：暂无

八、IBM Cognos

信誉等级：☆☆☆☆☆☆

主流产品：Cognos

产品评价：功能全、比较适合大中型企业，对于中小企业应用来讲有些大材小用。行业属性： 通用型

企业规模：中型或大型企业

产品参数：

Cognos Powerplay Transformation Server：

负责将数据源变成数据立方体;

Cognos Powerplay Enterprise Server：

负责将数据立方体以OLAP分析、OLAP报表等方式展现出来;

Cognos ReportNet Server：

负责实现基于数据库的数据查询、报表制作、仪表盘制作、报表/仪表盘展示等等;注意事项：

小型企业应用IBM COgnos有些大材小用，建议中型及中型以上企业应用。产品费用：比较昂贵，一般比较适合大型企业

**第二篇：酒店品牌**

酒店品牌

1.洲际国际酒店集团（Intercontinental Hotels Group）

旗下主要品牌： 洲际, 皇冠, 假日, 快捷

集团简介：洲际酒店集团是亚太地区及全球最大并拥有最多酒店品牌的酒店管理公司。在100多个国家里拥有、运营 及管理的酒店有3,500多家，酒店客房超过54.5万间。旗下的主要酒店品牌包括：洲际酒店及度假村，皇冠假日酒店及度 假村，假日酒店及度假村，快捷假日酒店。

在亚太地区，洲际酒店集团在23个国家经营160多家酒店。洲际酒店集团是一家英国公司，每年接待超过1.5亿客人，为旅客提供客房为主，并拥有有限的餐饮设施。

2．雅高集团（ACCOR）

旗下主要品牌： 索菲特、诺富特、美居酒店、雅高套房饭店、宜必思饭店、一级方程式汽车旅馆、红屋顶旅馆等

集团简介：法国雅高集团（ACCOR）是全球最大的酒店集团之一，法语意为和谐。雅高集团旗下共有14个品牌，其中7 个核心品牌，索菲特、诺富特、美居酒店、雅高套房饭店、宜必思饭店、一级方程式汽车旅馆、红屋顶旅馆等，以高品质 的设施和全面周到的服务获得极高声誉。

3．香格里拉酒店集团

旗下主要品牌：香格里拉酒店

集团简介：香格里拉酒店以香港为大本营，目前已是亚洲区最大的豪华酒店集团，且被视为世界最佳的酒店管理集团 之一。集团的47家酒店遍布亚洲和中东地区的主要城市以及大部分度假胜地。其中7家为商贸饭店，它是香格里拉酒店集 团的另一酒店品牌。

4．喜达屋国际酒店集团（Starwood Hotels &Resorts Worldwide）

旗下主要品牌： 瑞吉斯, 至尊精选, 喜来登, 威斯汀, 福朋及W 酒店，美丽殿

集团简介：喜达屋集团是全球最大的饭店及娱乐休闲集团之一。1998年，喜达屋收购了威斯汀饭店度假村国际集团（Westin Hotels&Resorts Worldwide,Inc)和它的几个分公司（包括威斯汀和威斯汀联合公司），还收购了美国国际电话电 报公司，并取名为Sheraton Holding Corporation。1999年10月，收购了维斯塔那（Vistana)股份有限公司（更名为喜达 屋度假所有权股份有限公司），这一切使得喜达屋集团在众多饭店及娱乐休闲集团中处于领先地位。喜达屋集团运用直接 或间接由其子公司管理的方法来经营饭店和娱乐休闲业务。

酒店设计公司 1.HBA/Hirsch Bedner Assocciates Design Consultants

HBA作为全球酒店室内设计业的领袖，把握着全球旅游者的脉搏，跟随着行业的潮流而动。很多国际知名的酒店的背后，都闪耀着HBA的设计师们无限的创造力。总部坐落于美国洛杉矶市的HBA，在全球拥有12 家分公司。HBA的设计哲学是：“在时间和预算允许的范围内，打造出一个精心策划的，把梦境，剧场，及舒适融为一体的空间感。2.Bilkey Llinas Design(BLD)

Bilkey Llinas Design（BLD）总部位于美国佛罗里达，公司的目标是致力于寻

找设计变化要求的根本的解决方案，并不断研究最前沿的行业，而且有专业的建筑师和设计师的公司团队，努力创新的概念，不仅能够吸引商务旅行者也满足旅客的要求。3.CHENG CHUNG DESIGN

（CCD）香港郑中设计事务所（以下简称CCD）系由香港著名设计师JOE CHENG——郑忠先生创立，专业为国际品牌酒店提供室内设计及顾问服务，是国际顶级品牌酒店室内设计机构之一。CCD的国际化的团队及专业技术令其与时俱进，在行业内始终能够与时俱进，保持前瞻性的创新。CCD获得了数十个国际室内设计奖项，金钥匙奖，IIDA，HD等。CCD在美国《室内设计》杂志2025年10月的全球酒店室内设计百大排名中名列第4，是唯一进入该排名前75名的亚洲设计公司。作为中国境内最庞大及最具富有成效的设计事务所之一，CCD引领了行业的发展。CCD自创立至今，先后为10个国际酒店管理集团，20个国际品牌，为亚洲地区的100家豪华酒店成功提供了的室内设计及顾问服务工作。CCD先后为国际品牌酒店管理集团，如Hilton（希尔顿）, Marriott（万豪）, Accor（雅高）Wyndham（温德姆）,InterContinental Hotels Group（洲际）, Starwood（喜达屋）,Raffles（莱佛士）,朗豪酒店（Langham）等成功服务，并获得了他们的高度认可。CCD为全球客户提供综合的一体化设计服务和方案，以前沿的设计理念，最新的技术工艺及独一无二的产品形象等提升服务项目的产业价值。CCD将为中国酒店设计事业走向世界而继续努力。

4.姜峰室内设计有限公司 JIANG&ASSOCIATES INTERIOR DESIGN CO.,LTD

酒店设计和购物中心是J&AD的重点设计领域。通过近几年完成的一系列颇具

影响力的国际知名品牌酒店的设计，J&A已经探索出一条风格鲜明、功能与艺术并重，且符合国际发展潮流的设计路线。

如何入手规划 设计是把一种计划、规划、设想通过视觉的形式传达出来的活动过程。人类通过劳动改造世界，创造文明，创造物质财富和精神财富，而最基础、最主要的创造活动是造物。设计便是造物活动进行预先的计划，可以把任何造物活动的计划技术和计划过程理解为设计。是对事物一个统筹规划的一个过程的总称。

对于酒店装修，不但要求装修工程合格，还要求酒店的装修风格独特，只有这样才能够吸引更多的客户，还有就是酒店的施工质量一定要高： 1.对于墙面的处理。在国内一般采用石材、瓷砖的比较多，在国外外墙主要采用涂料。对于快捷酒店装修来说，如果是新建的话，那么酒店的设计一定要做好，如果仅仅是改造，那么就不必打掉瓷砖和石材，可以直接在上面弄一层腻子。2.酒店灯光的设置。好的灯光设计应当与建筑融为一体。现在灯的品种有许多种，国内与跑马灯配套的灯泡寿命欠佳功率也过大，可用调低电压的办法解决。国外的招牌字体一般是通体发亮，技术比国内的要高。3.涂料的选择。涂料分为水性、油性和石性漆几种。真石漆是一种高档涂料，价格在 100 多元 一平方米。如果选择普通涂料，最好选择亚光。要注意的是一定要选择知名品牌和由专业施工单位来做。

4.提高投入效果比。一个好的酒店设计方案，应当是投入效果比最大化，也就是用最少的钱装修出最好的效果，所以在前期，设计方案是非常重要的。

**第三篇：BI学习心得**

BI学习心得 BI是什么

通过学习理解，可将BI归结为一句话：从数据中挖取利润。

1．1 初识BI 初识BI的概念，是从《BI相关知识介绍.ppt》开始的。通读之后，我了解三件事： 1）、沃尔玛将啤酒和尿布摆在一起出售，提高了啤酒的销量； 2）、BI是Business Intelligence的缩写，商务智能。官方定义：是一种运用了数据仓库、在线分析和数据挖掘等技术来处理和分析数据的崭新技术。3）、BI的分析过程图，如下：

图1.1—BI Analyse的过程

同时，幻灯片中还介绍了大量的中英文概念，让我非常的混乱。为了更好的学习了解BI，首先需要克服的就是名词解释，其次是通过实例了解BI的应用。1．2 常见名词浅释

 ETL：数据抽取（Extract）、转换（Transform）、清洗（Cleansing）、装载（Load）的过程。构建数据仓库的重要一环，用户从数据源抽取出所需的数据，经过数据清洗，最终按照预先定义好的数据仓库模型，将数据加载到数据仓库中去。 Data Warehouse：数据仓库，是在企业管理和决策中面向主题的、集成的、与时间相关的、不可修改的数据集合。与其他数据库应用不同的是，数据仓库更像一种过程，对分布在企业内部各处的业务数据的整合、加工和分析的过程。而不是一种可以购买的产品。 Data mart：数据集市，或者叫做“小数据仓库”。如果说数据仓库是建立在企业级的数据模型之上的话。那么数据集市就是企业级数据仓库的一个子集，他主要面向部门级业务，并且只是面向某个特定的主题。数据集市可以在一定程度上缓解访问数据仓库的瓶颈。 OLAP（On-Line Analytical Processing）：联机在线分析处理，是一类软件技术，它们使用户（业务分析师、经理和执行官）能够以交互形式快速、一致地探查数据，用户看到的是经过转换后的原始数据的各种信息视图，它们可以反映业务的真实维数。 ROLAP：基于Codd的12条准则，各个软件开发厂家见仁见智，其中一个流派，认为可以沿用关系型数据库来存储多维数据，于是，基于稀疏矩阵表示方法的星型结构（star schema）就出现了。后来又演化出雪花结构。为了与多维数据库相区别，则把基于关系型数据库的OLAP称为Relational OLAP，简称ROLAP。代表产品有Informix Metacube、Microsoft SQL Server OLAP Services。 MOLAP：Arbor Software严格遵照Codd的定义，自行建立了多维数据库，来存放联机分析系统数据，开创了多维数据存储的先河，后来的很多家公司纷纷采用多维数据存储。被人们称为Muiltdimension OLAP，简称MOLAP，代表产品有Hyperion(原Arbor Software)Essbase、Showcase Strategy等。 Client OLAP：相对于Server OLAP而言。部分分析工具厂家建议把部分数据下载到本地，为用户提供本地的多维分析。代表产品有Brio Designer，Business Object。 DSS：决策支持系统（Decision Support System），相当于基于数据仓库的应用。决策支持就是在收集所有有关数据和信息，经过加工整理，来为企业决策管理层提供信息，为决策者的决策提供依据。 Ad hoc query：即席查询，数据库应用最普遍的一种查询，利用数据仓库技术，可以让用户随时可以面对数据库，获取所希望的数据。 EIS：主管信息系统（Executive Information System），指为了满足无法专注于计算机技术的领导人员的信息查询需求，而特意制定的以简单的图形界面访问数据仓库的一种应用。

 BPR：业务流程重整（Business Process Reengineering），指利用数据仓库技术，发现并纠正企业业务流程中的弊端的一项工作，数据仓库的重要作用之一。 BI：商业智能（Business Intelligence），指数据仓库相关技术与应用的通称。指利用各种智能技术，来提升企业的商业竞争力。 Data Mining：数据挖掘，Data Mining是一种决策支持过程，它主要基于AI、机器学习、统计学等技术，高度自动化地分析企业原有的数据，做出归纳性的推理，从中挖掘出潜在的模式，预测客户的行为，帮助企业的决策者调整市场策略，减少风险，做出正确的决策  CRM：客户关系管理（Customer Relationship Management），数据仓库是以数据库技术为基础但又与传统的数据库应用有着本质区别的新技术，CRM就是基于数据仓库技术的一种新应用。但是，从商业运作的角度来讲，CRM其实应该算是一个古老的“应用”了。比如，酒店对客人信息的管理，如果某个客人是某酒店的老主顾，那么该酒店很自然地会知道这位客人的某些习惯和喜好，如是否喜欢靠路边，是否吸烟，是否喜欢大床，喜欢什么样的早餐，等等。当客人再次光临时，不用客人自己提出来，酒店就会提供客人所喜欢的房间和服务。这就是一种CRM。 Meta Data：元数据，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。

1．3 实例应用

BI应用实例：

某家上市公司的销售副总裁每周五下午要和全国各大区的销售总监开视频会议，一般情况下他需要事先了解一下公司上周的销售情况，这时他进入公司的商业智能系统，查看销售周报表，报表的横坐标是时间，纵坐标有：“销售目标”、“实际收入”、“收入指标达成率”。这张报表显示了每周销售变化情况。

副总裁为了进一步弄清楚某一大区及大区下属的几大城市的销售数据，只要使用钻取功能，用鼠标轻轻在屏幕上点击几下，该大区及大区下各城市的销售数据便一目了然。钻取即这种从“全国水平”（宏观）到“地区水平”（微观）的分析过程。然后，副总裁需要了解在某一区域的销售数据中，哪些是直销完成的，哪些是渠道完成的，这种从报表视图的某一个角度到另外一个角度的视角转换过程被称为“切片”。钻取和切片的操作都非常简便，只需用鼠标简单地在电子报表上轻点几下即可完成。在上面的分析过程中，涉及到了“维”这个概念，即观察问题的角度。

这位销售副总裁的分析维度包括：时间、大区、销售模式，以这三个维为三个轴线来观察数据，就形成了一个信息立方体，立方体的每一个小单元都包含了企业管理者需要追踪的消息，包括一个特定地区、特定时间、特定销售渠道的各项指标的全部信息，通过对立方体的切片或切块分割，可以从不同的视角观察各种数据。通过“钻取”和“切片”，这位副总裁知道了销售情况的变化是哪个大区、哪些城市、直销或者代理渠道出现了问题，这种灵活快速的查询和分析如果通过过去的ERP报表查询是很难完成的。

通过实例可知，使用BI系统，可以将不同的数据源关联对应在一起，对企业信息进行管理、分析、自动生成报表，提高分析决策的效率和质量。

1．4 BI的定义

商务智能是指将企业的各种数据及时地转换为企业管理者感兴趣的信息（或者知识），并以各种方式展现出来，帮助企业管理者进行科学决策，加强企业的竞争优势。这里的数据不仅仅指企业内部的各种数据，而且包括企业外部的数据，例如行业、市场状况和客户资源的数据等等。

从技术角度看，商务智能的过程是企业的决策人员以企业中的数据仓库为基础，经由联机分析处理工具、数据挖掘工具加上决策规划人员的专业知识，从数据中获得有用的信息和知识，帮助企业获取利润。

从应用角度看，商务智能帮助用户对商业数据进行联机分析处理和数据挖掘，例如预测发展趋势、辅助决策、对客户进行分类、挖掘潜在客户等等。

从数据角度看，商务智能使得很多事务性的数据经过抽取、转换之后存入数据仓库，经过聚集、切片或者分类等操作之后形成有用的信息、规则，来帮助企业的决策者进行正确的决策。

综上所述，我认为

BI的定义：商务智能是利用各种技术，对数据进行收集、存储、分析并提供可视化界面显示，从而更容易、快捷的获得信息知识，帮助管理者做出更好的业务决策。BI技术简析

通过阅读资料，我认为把商业智能看成是一种解决方案的说法比较恰当。它是将数据仓库、OLAP和数据挖掘等技术组合到一起的综合运用。BI技术的关键是将来自不同的业务系统数据库中的数据，提取出来，进行清理以保证数据的正确性，然后经过抽取（Extraction）、转换（Transformation）和装载（Load），即ETL过程，合并到数据仓库里，从而得到企业数据的一个全局视图，在此基础上利用合适的查询和分析工具、数据挖掘工具、OLAP工具等对其进行分析和处理（这时信息变为辅助决策的知识），最后将知识呈现给管理者，为管理者的决策过程提供支持。商业智能的这个基本过程如图所示：

图2—BI结构图

一些文章中将BI分为四个主要阶段：数据预处理、建立数据仓库、数据分析、数据展现。同时认为数据仓库、OLAP和数据挖掘技术是商业智能的三大技术支柱。2．1 数据预处理——ETL ETL是商业智能（BI）的重要组成部分，是英文Extract-Transform-Load的缩写，即填充、更新数据仓库的数据抽取、转换、装载的过程。这是实现商业智能之前的数据采集步骤。这一步骤完成之后，对库中数据的数据挖掘、分析处理才可以进行。

对于ETL而言，“是什么”是很容易理解的，也就是将分散的、不易利用的数据进行整理，变成规则清晰的、易于利用的、（可能同时还是）集中的数据。

2．1．1 ETL设计

ETL的设计分三部分：数据抽取、数据的清洗转换、数据的加载。数据的抽取是从各个不同的数据源抽取到ODS(OperationalDataStore，操作型数据存储)中——这个过程也可以做一些数据的清洗和转换)，在抽取的过程中需要挑选不同的抽取方法，尽可能的提高ETL的运行效率。ETL三个部分中，花费时间最长的是“T”(Transform，清洗、转换)的部分，一般情况下这部分工作量是整个ETL的2/3。数据的加载一般在数据清洗完了之后直接写入DW(DataWarehousing，数据仓库)中去。

2．1．2 ETL实现

ETL的实现有多种方法，常用的有三种。一种是借助ETL工具(如Oracle的OWB、SQLServer2000的DTS、SQLServer2005的SSIS服务、Informatic等)实现，一种是SQL方式实现，另外一种是ETL工具和SQL相结合。前两种方法各有各的优缺点，借助工具可以快速的建立起ETL工程，屏蔽了复杂的编码任务，提高了速度，降低了难度，但是缺少灵活性。SQL的方法优点是灵活，提高ETL运行效率，但是编码复杂，对技术要求比较高。第三种是综合了前面二种的优点，会极大地提高ETL的开发速度和效率。

2．1．3 数据的抽取

这一部分需要在调研阶段做大量的工作，首先要搞清楚数据是从几个业务系统中来,各个业务系统的数据库服务器运行什么DBMS,是否存在手工数据，手工数据量有多大，是否存在非结构化的数据等等，当收集完这些信息之后才可以进行数据抽取的设计。

1、对于与存放DW的数据库系统相同的数据源处理方法

这一类数据源在设计上比较容易。一般情况下，DBMS(SQLServer、Oracle)都会提供数据库链接功能，在DW数据库服务器和原业务系统之间建立直接的链接关系就可以写Select语句直接访问。

2、对于与DW数据库系统不同的数据源的处理方法

对于这一类数据源，一般情况下也可以通过ODBC的方式建立数据库链接——如SQLServer和Oracle之间。如果不能建立数据库链接，可以有两种方式完成，一种是通过工具将源数据导出成.txt或者是.xls文件，然后再将这些源系统文件导入到ODS中。另外一种方法是通过程序接口来完成。

3、对于文件类型数据源(.txt,.xls)可以培训业务人员利用数据库工具将这些数据导入到指定的数据库，然后从指定的数据库中抽取。或者还可以借助工具实现，如SQLServer2005的SSIS服务的平面数据源和平面目标等组件导入ODS中去。

4、增量更新的问题

对于数据量大的系统，必须考虑增量抽取。一般情况下，业务系统会记录业务发生的时间，我们可以用来做增量的标志,每次抽取之前首先判断ODS中记录最大的时间，然后根据这个时间去业务系统取大于这个时间所有的记录。利用业务系统的时间戳，一般情况下，业务系统没有或者部分有时间戳。

2．1．4 数据的清洗转换

一般情况下，数据仓库分为ODS、DW两部分。通常的做法是从业务系统到ODS做清洗，将脏数据和不完整数据过滤掉，在从ODS到DW的过程中转换，进行一些业务规则的计算和聚合。

1、数据清洗

数据清洗的任务是过滤那些不符合要求的数据，将过滤的结果交给业务主管部门，确认是否过滤掉还是由业务单位修正之后再进行抽取。不符合要求的数据主要是有不完整的数据、错误的数据、重复的数据三大类。

(1)不完整的数据：这一类数据主要是一些应该有的信息缺失，如供应商的名称、分公司的名称、客户的区域信息缺失、业务系统中主表与明细表不能匹配等。对于这一类数据过滤出来，按缺失的内容分别写入不同Excel文件向客户提交，要求在规定的时间内补全。补全后才写入数据仓库。

(2)错误的数据：这一类错误产生的原因是业务系统不够健全，在接收输入后没有进行判断直接写入后台数据库造成的，比如数值数据输成全角数字字符、字符串数据后面有一个回车操作、日期格式不正确、日期越界等。这一类数据也要分类，对于类似于全角字符、数据前后有不可见字符的问题，只能通过写SQL语句的方式找出来，然后要求客户在业务系统修正之后抽取。日期格式不正确的或者是日期越界的这一类错误会导致ETL运行失败，这一类错误需要去业务系统数据库用SQL的方式挑出来，交给业务主管部门要求限期修正，修正之后再抽取。

(3)重复的数据：对于这一类数据——特别是维表中会出现这种情况——将重复数据记录的所有字段导出来，让客户确认并整理。

数据清洗是一个反复的过程，不可能在几天内完成，只有不断的发现问题，解决问题。对于是否过滤，是否修正一般要求客户确认，对于过滤掉的数据，写入Excel文件或者将过滤数据写入数据表，在ETL开发的初期可以每天向业务单位发送过滤数据的邮件，促使他们尽快地修正错误,同时也可以做为将来验证数据的依据。数据清洗需要注意的是不要将有用的数据过滤掉，对于每个过滤规则认真进行验证，并要用户确认。

2、数据转换 数据转换的任务主要进行不一致的数据转换、数据粒度的转换，以及一些商务规则的计算。

(1)不一致数据转换：这个过程是一个整合的过程，将不同业务系统的相同类型的数据统一，比如同一个供应商在结算系统的编码是XX0001,而在CRM中编码是YY0001，这样在抽取过来之后统一转换成一个编码。

(2)数据粒度的转换：业务系统一般存储非常明细的数据，而数据仓库中数据是用来分析的，不需要非常明细的数据。一般情况下，会将业务系统数据按照数据仓库粒度进行聚合。

(3)商务规则的计算：不同的企业有不同的业务规则、不同的数据指标，这些指标有的时候不是简单的加加减减就能完成，这个时候需要在ETL中将这些数据指标计算好了之后存储在数据仓库中，以供分析使用。

2．1．5 数据的装载

ETL的装载部分，可查询的资料较少。仅查阅到与抽取在一起的部分描述内容。ETL的核心还是在于T这个过程，也就是转换，而抽取和装载一般可以作为转换的输入和输出，或者，它们作为一个单独的部件，其复杂度没有转换部件高。和OLTP系统中不同，那里充满这单条记录的insert、update和select等操作，ETL过程一般都是批量操作，例如它的装载多采用批量装载工具，一般都是DBMS系统自身附带的工具，例如Oracle SQLLoader和DB2的autoloader等。

2．1．6 ETL日志、警告发送

1、ETL日志

ETL日志分为三类。一类是执行过程日志，这一部分日志是在ETL执行过程中每执行一步的记录，记录每次运行每一步骤的起始时间，影响了多少行数据，流水账形式。一类是错误日志，当某个模块出错的时候写错误日志，记录每次出错的时间、出错的模块以及出错的信息等。第三类日志是总体日志，只记录ETL开始时间、结束时间是否成功信息。如果使用ETL工具,ETL工具会自动产生一些日志，这一类日志也可以作为ETL日志的一部分。记录日志的目的是随时可以知道ETL运行情况，如果出错了，可以知道哪里出错。

2、警告发送

如果ETL出错了，不仅要形成ETL出错日志，而且要向系统管理员发送警告。发送警告的方式多种，一般常用的就是给系统管理员发送邮件，并附上出错的信息，方便管理员排查错误。

ETL是BI项目的关键部分，也是一个长期的过程，只有不断的发现问题并解决问题，才能使ETL运行效率更高，为BI项目后期开发提供准确的数据。

2．2 BI支撑技术——数据仓库/数据集市

数据仓库：商业智能的基础。

对于一个企业来说，最关键也最为重要的是，如何以一种有效的方式逐步整理各个业务处理系统中积累下来的历史数据，并通过灵活有效的方式为各级业务人员提供统一的信息视图，从而在整个企业内实现真正的信息共享。数据仓库技术正好满足了这一需求。数据仓库是商业智能系统的基础，如果没有数据仓库，没有企业数据的融合，数据分析就成为了无源之水。

数据仓库定义为：“一个面向主题的、集成的、反映历史变化的、相对稳定的数据集合，用于支持管理决策”。与传统的事物处理系统有不同的是，数据仓库是一种过程，它是对分布在组织或企业内部各处的业务数据进行整合、加工和分析的过程。而不是一种可以购买的产品。

2．2．1 数据仓库常用术语

 数据仓库

数据仓库是一个支持管理决策的数据集合。数据是面向主题的、集成的、不易丢失的并且是时间变量。

数据仓库是所有操作环境和外部数据源的快照集合。它并不需要非常精确，因为它必须在特定的时间基础上从操作环境中提取出来。 数据集市

数据仓库只限于单个主题的区域，例如顾客、部门、地点等。数据集市在从数据仓库获取数据时可以依赖于数据仓库，或者当它们从操作系统中获取数据时就不依赖于数据仓库。 事实

事实是数据仓库中的信息单元，也是多维空间中的一个单元，受分析单元的限制。事实存储于一张表中（当使用关系数据库时）或者是多维数据库中的一个单元。每个事实包括关于事实（收入、价值、满意记录等）的基本信息，并且与维度相关。在某些情况下，当所有的必要信息都存储于维度中时，单纯的事实出现就是对于数据仓库足够的信息。我们稍后讨论有关缺无事实的情况。 维度

维度是绑定由坐标系定义的空间的坐标系的轴线。数据仓库中的坐标系定义了数据单元，其中包含事实。

坐标系的一个例子就是带有 x 维度和 y 维度的 Cartesian（笛卡尔）坐标系。在数据仓库中，时间总是维度之一。 数据挖掘

在数据仓库的数据中发现新信息的过程被称为数据挖掘，这些新信息不会从操作系统中获得。 分析空间

分析空间是数据仓库中一定量的数据，用于进行数据挖掘以发现新信息同时支持管理决策。 切片

一种用来在数据仓库中将一个维度中的分析空间限制为数据子集的技术。 切块

一种用来在数据仓库中将多个维度中的分析空间限制为数据子集的技术。 星型模式 一种使用关系数据库实现多维分析空间的模式，称为星型模式。星型模式将在本白皮书中稍后进行进一步讨论。 雪花模式

不管什么原因，当星型模式的维度需要进行规范化时，星型模式就演进为雪花模式。

2．2．2 面向主题（Subject-oriented）

“主题”是一个较为抽象的概念，是指用户使用数据仓库进行决策时所关心的重点方面。从信息管理的角度看，主题是在一个较高的管理层次上对数据进行综合、归类所形成的分析对象；从数据组织的角度看，主题就是一些数据集合，这些数据集合对分析对象作了比较完整的、一致的描述，这种描述不仅涉及到数据本身，还涉及到数据之间的关系。

“面向主题”则表明了数据仓库中数据组织的基本原则，是指数据仓库内的信息是按主题进行组织的，而不像传统事物处理系统那样单一地按照业务功能及性能要求进行组织。传统的操作型系统是围绕公司的应用进行组织。如对一个电信公司来说，应用问题可能是营业受理、专业计费和客户服务等，而主题范围可能是客户、套餐、缴费和欠费等。

2．2．3 集成（Integrated）

“集成”是指数据仓库中的信息并不是对各个数据源简单的选择、抽取，而是首先进行一系列的加工、整理和转换等来消除源数据中的不一致；同时按照本行业的逻辑模型设计便于查询及分析的数据仓库。然后按照组织或企业的需求，针对不同的主题对数据进行某种程度的综合、概括和聚集，将源数据加载进数据仓库。经过这样的处理，数据就具有了集成性，可以用于决策分析。

2．2．4 反映历史变化（Time-variant）

“反映历史变化”是指数据仓库内的信息并不只是反映企业当前的状态，而是记录了从过去某一时点到当前各个阶段的信息。通过这些信息，可以对企业的发展历程和未来趋势做出定量分析和预测。而信息本身相对稳定，是指一旦某个数据进入数据仓库以后，一般很少进行修改，更多的是对信息进行查询操作。

2．2．5 相对稳定（Non-volatile）

“相对稳定”是指数据一旦进入数据仓库，一般情况下会被长期保留，所涉及的数据操作也主要是查询、分析，很少会被修改或删除，通常也只需要定期地加载和刷新。相对稳定性保证了数据仓库中的数据能够真实地反映历史变化。

数据仓库中所存放的数据在物理上是分离的。由于这种分离，数据仓库不需要事务处理、恢复和并发控制机制。通常它只需要两种数据访问：数据的初始化装入和数据访问。

2．2．6 了解实时数据仓库的定义

实时数据仓库是两种事物的组合：实时行为和数据仓库。实时行为是一种即时发生的行为。行为可以是任何事情，如超市中小商品的销售行为。一旦行为完成，就有关于它的数据。数据仓库捕获有关商业行为的数据，而实时数据仓库在商业行为发生时就捕获数据。当商业行为完成时，相关数据就已经进入到数据仓库并且能立即使用。换句话说，实时数据仓库是这样一个系统，只要行为发生、数据变得可用时，就能从中获得信息。

实时vs.近似实时之间的区别可以下面一个词概括：反应时间。反应时间是指行为的完成时间和该行为的数据在数据仓库中可利用时间之间的延迟。实时情况下，反应时间是可以忽略不计的；在近似实时情况下，反应时间是一段时间，比如2小时。为了使系统变得更精确，一些人使用了“实时”这个概念。

2．3 BI分析技术——OLAP 2．3．1 什么是OLAP

当今的数据处理大致可以分成两大类：联机事务处理OLTP（on-line transaction processing）、联机分析处理OLAP（On-Line Analytical Processing）。OLTP是传统的关系型数据库的主要应用，主要是基本的、日常的事务处理，例如银行交易。OLAP是数据仓库系统的主要应用，支持复杂的分析操作，侧重决策支持，并且提供直观易懂的查询结果。下表列出了OLTP与OLAP之间的比较。

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是使分析人员、管理人员或执行人员能够从多角度对信息进行快速、一致、交互地存取,从而获得对数据的更深入了解的一类软件技术。OLAP的目标是满足决策支持或者满足在多维环境下特定的查询和报表需求,它的技术核心是“维”这个概念。

区别于传统的联机事务处理(OLTP)系统，OLAP有12条准则：

1、OLAP模型必须提供多维概念视图

2、透明性准则

3、存取能力推测

4、稳定的报表能力

5、客户/服务器体系结构

6、维的等同性准则

7、动态的稀疏矩阵处理准则

8、多用户支持能力准则

9、非受限的跨维操作

10、直观的数据操纵

11、灵活的报表生成

12、不受限的维与聚集层次 虽然随着技术的发展，部分准则有所突破，但这些准则仍然是OLAP技术的基础。

2．3．2 OLAP核心技术——维

“维”是人们观察客观世界的角度,是一种高层次的类型划分。“维”一般包含着层次关系,这种层次关系有时会相当复杂。通过把一个实体的多项重要的属性定义为多个维(dimension)，使用户能对不同维上的数据进行比较。因此OLAP也可以说是多维数据分析工具的集合。

OLAP的基本多维分析操作有钻取（roll up和drill down）、切片（slice）和切块（dice）、以及旋转（pivot）、drill across、drill through等。 钻取是改变维的层次，变换分析的粒度。它包括向上钻取（roll up）和向下钻取（drill down）。♢ roll up是在某一维上将低层次的细节数据概括到高层次的汇总数据，或者减少维数； ♢ 而drill down则相反，它从汇总数据深入到细节数据进行观察或增加新维。 切片和切块是在一部分维上选定值后，关心度量数据在剩余维上的分布。如果剩余的维只有两个，则是切片；如果有三个，则是切块。 旋转是变换维的方向，即在表格中重新安排维的放置（例如行列互换）。 Drill across：（未找到相关介绍） Drill through：（未找到相关介绍）

2．3．3 OLAP的分类

OLAP有多种实现方法，根据存储数据的方式不同可以分为ROLAP、MOLAP、HOLAP。ROLAP 表示基于关系数据库的OLAP实现（Relational OLAP）。以关系数据库为核心,以关系型结构进行多维数据的表示和存储。ROLAP将多维数据库的多维结构划分为两类表:一类是事实表,用来存储数据和维关键字;另一类是维表,即对每个维至少使用一个表来存放维的层次、成员类别等维的描述信息。维表和事实表通过主关键字和外关键字联系在一起,形成了“星型模式”。对于层次复杂的维,为避免冗余数据占用过大的存储空间,可以使用多个表来描述,这种星型模式的扩展称为“雪花模式”。ROLAP通常使用三层架构，其数据库层和应用逻辑层是分离的，并且其数据存储在关系型数据库中

MOLAP 表示基于多维数据组织的OLAP实现（Multidimensional OLAP）。以多维数据组织方式为核心,也就是说,MOLAP使用多维数组存储数据。多维数据在存储中将形成“立方块（Cube）”的结构,在MOLAP 中对“立方块”的“旋转”、“切块”、“切片”是产生多维数据报表的主要技术。

MOLAPS和ROLAPS相比，具有更短的处理时间，更快的响应速度，并且对于用户需求更为灵活。而ROLAPS在处理的数据总量和支持的客户数方面强于MOLAPS。

Hybrid online analytical processing(混合联机分析处理，HOLAP)，HOLAP表示基于混合数据组织的OLAP实现。如低层是关系型的，高层是多维矩阵型的。这种方式具有更好的灵活性。从字面上我们可以大致看出，HOLAP是上面两种分类的融合，它同时利用了多维数据库和关系数据库。HOLAP的优势在于它可以在性能需求和存储量需求间找到一个平衡点，从而充分利用系统潜能。

还有其他的一些实现OLAP的方法，如提供一个专用的SQL Server，对某些存储模式（如星型、雪片型）提供对SQL查询的特殊支持。

2．3．4 OLAP工具描述

OLAP 工具是针对特定问题的联机数据访问与分析。它通过多维的方式对数据进行分析、查询和报表。维是人们观察数据的特定角度。例如，一个企业在考虑产品的销售情况时，通常从时间、地区和产品的不同角度来深入观察产品的销售情况。这里的时间、地区和产品就是维。而这些维的不同组合和所考察的度量指标构成的多维数组则是OLAP分析的基础，可形式化表示为（维1，维2，„„，维n，度量指标），如（地区、时间、产品、销售额）。多维分析是指对以多维形式组织起来的数据采取切片（Slice）、切块（Dice）、钻取（Drill-down和Roll-up）、旋转（Pivot）等各种分析动作，以求剖析数据，使用户能从多个角度、多侧面地观察数据库中的数据，从而深入理解包含在数据中的信息。

2．4 BI分析技术——数据挖掘

2．4．1 什么是数据挖掘

数据挖掘又称知识发现（Knowledge Discovery in Database，简称KDD），是从大量数据中抽取有意义的、隐含的、以前未知的并有潜在使用价值的知识的过程。数据挖掘是一个多学科交叉性学科，它涉及统计学、机器学习、数据库、模式识别、可视化以及高性能计算等多个学科。利用数据挖掘技术可以分析各种类型的数据，例如结构化数据、半结构化数据以及非结构化数据、静态的历史数据和动态数据流数据等。

2．4．2 数据挖掘分类

数据挖掘涉及的学科领域和方法很多，有人工智能、数据统计、可视化、并行计算等。数据挖掘有多种分类方法。

 根据挖掘任务

可分为分类模型发现、聚类、关联规则发现、序列分析、偏差分析、数据可视化等。(1)分类(Classification)其旨在生成一个分类函数或分类模型，该模型能把数据库中的数据项映射到给定类别中的某一个。既可以用此模型分析已有的数据，也可以用它来预测未来的数据。

(2)聚集(Clustering)聚集是对记录分组，把相似的记录在一个聚集里。聚集和分类的区别是聚集不依赖于预先定义好的类，不需要训练集。

(3)数据可视化(Description and Visualization)数据可视化严格地讲不是一个单独的数据挖掘任务，它被用来支持其他挖掘任务。可视化是采用图形、图表等易于理解的方式表达数据挖掘结果。

(4)关联规则(Affinity grouping or association rules)关联规则是寻找数据库中值的相关性，主要是寻找在同一个事件中出现的不同项的相关性，比如在一次购买活动中所买不同商品的相关性。(5)序列分析(Sequence Analysis)序列模式分析同样也是试图找出数据之间的联系。但它的侧重点在于分析数据之间前后(因果)关系，因此对数据往往要求引入时间属性。序列模式分析非常适于寻找事物的发生趋势或重复性模式。

(6)偏差分析(Deviation Analysis)是用来发现与正常情况不同的异常和变化，并进一步分析这种变化是否是有意的诈骗行为，还是正常的变化。如果是异常行为，则提示预防措施;如果是正常的变化，那么就需要更新数据库记录。

 根据挖掘对象

可分为关系数据库、面向对象数据库、空间数据库、时态数据库、文本数据源、多媒体数据库、异质数据库以及环球网Web。

 根据挖掘方法

可粗略地分为：机器学习方法、统计方法、神经网络方法、决策树、可视化、最近邻技术等。在机器学习中，可细分为归纳学习方法(决策树、规则归纳等)、基于范例学习、遗传算法等;在统计方法中，可细分为:回归分析(多元回归、自回归等)、判别分析(贝叶斯判别、费歇尔判别、非参数判别等)、聚类分析(系统聚类、动态聚类等)、探索性分析(主元分析法、相关分析法等)等。

2．4．3 数据挖掘技术

数据挖掘技术可分为描述型数据挖掘和预测型数据挖掘两种。描述型数据挖掘包括数据总结、聚类及关联分析等。预测型数据挖掘包括分类、回归及时间序列分析等。

1、数据总结

继承于数据分析中的统计分析。数据总结目的是对数据进行浓缩,给出它的紧凑描述。传统统计方法如求和值、平均值、方差值等都是有效方法。另外还可以用直方图、饼状图等图形方式表示这些值。广义上讲，多维分析也可以归入这一类。

2、聚类分析

是把整个数据库分成不同的群组。它的目的是使群与群之间差别很明显，而同一个群之间的数据尽量相似。这种方法通常用于客户细分。在开始细分之前不知道要把用户分成几类，因此通过聚类分析可以找出客户特性相似的群体，如客户消费特性相似或年龄特性相似等。在此基础上可以制定一些针对不同客户群体的营销方案。

聚类分析是根据物以类聚的原理，将本身没有类别的样本聚集成不同的组，并且对每一个这样的组进行描述的过程。其主要依据是聚到同一个组中的样本应该彼此相似，而属于不同组的样本应该足够不相似。

以客户关系管理为例，利用聚类技术，根据客户的个人特征以及消费数据，可以将客户群体进行细分。例如，可以得到这样的一个消费群体：女性占91%，全部无子女、年龄在31到40岁占70%，高消费级别的占64%，买过针织品的占91%，买过厨房用品的占89%，买过园艺用品的占79%。针对不同的客户群，可以实施不同的营销和服务方式，从而提高客户的满意度。

对于空间数据，根据地理位置以及障碍物的存在情况可以自动进行区域划分。例如，根据分布在不同地理位置的ATM机的情况将居民进行区域划分，根据这一信息，可以有效地进行ATM机的设置规划，避免浪费，同时也避免失掉每一个商机。

对于文本数据，利用聚类技术可以根据文档的内容自动划分类别，从而便于文本的检索。

3、关联分析

关联分析主要用于发现不同事件之间的关联性，即一个事件发生的同时，另一个事件也经常发生。关联分析的重点在于快速发现那些有实用价值的关联发生的事件。其主要依据是事件发生的概率和条件概率应该符合一定的统计意义。

对于结构化的数据，以客户的购买习惯数据为例，利用关联分析，可以发现客户的关联购买需要。例如，一个开设储蓄账户的客户很可能同时进行债券交易和股票交易，购买纸尿裤的男顾客经常同时购买啤酒等。利用这种知识可以采取积极的营销策略，扩展客户购买的产品范围，吸引更多的客户。通过调整商品的布局便于顾客买到经常同时购买的商品，或者通过降低一种商品的价格来促进另一种商品的销售等。

对于非结构化的数据，以空间数据为例，利用关联分析，可以发现地理位置的关联性。例如，85%的靠近高速公路的大城镇与水相邻，或者发现通常与高尔夫球场相邻的对象等。

4、分类分析

目的是构造一个分类函数或分类模型(也常常称作分类器)，该模型能把数据库中的数据项映射到给定类别中的某一个。要构造分类器，需要有一个训练样本数据集作为输入。训练集由一组数据库记录或元组构成,每个元组是一个由有关字段(又称属性或特征)值组成的特征向量,此外,训练样本还有一个类别标记。一个具体样本的形式可表示为:(v1, v2,...，vn；c)，其中vi表示字段值，c表示类别。

分类分析通过分析具有类别的样本的特点，得到决定样本属于各种类别的规则或方法。利用这些规则和方法对未知类别的样本分类时应该具有一定的准确度。其主要方法有基于统计学的贝叶斯方法、神经网络方法、决策树方法以及support vector machines等。利用分类技术，可以根据顾客的消费水平和基本特征对顾客进行分类，找出对商家有较大利益贡献的重要客户的特征，通过对其进行个性化服务，提高他们的忠诚度。

利用分类技术，可以将大量的半结构化的文本数据，如WEB页面、电子邮件等进行分类。可以将图片进行分类，例如，根据已有图片的特点和类别，可以判定一幅图片属于何种类型的规则。对于空间数据，也可以进行分类分析，例如，可以根据房屋的地理位置决定房屋的档次。

5、回归

是通过具有已知值的变量来预测其它变量的值。一般情况下，回归采用的是线性回归、非线性回归这样的标准统计技术。一般同一个模型既可用于回归也可用于分类。常见的算法有逻辑回归、决策树、神经网络等。

6、时间序列

时间序列是用变量过去的值来预测未来的值。

时间序列分析的是随时间而变化的事件序列，目的是预测未来发展趋势，或者寻找相似发展模式或者是发现周期性发展规律。

7、序列分析

序列分析技术主要用于发现一定时间间隔内接连发生的事件。这些事件构成一个序列，发现的序列应该具有普遍意义，其依据除了统计上的概率之外，还要加上时间的约束。

8、预测

预测与分类类似，但预测是根据样本的已知特征估算某个连续类型的变量的取值的过程，而分类则只是用于判别样本所属的离散类别而已。预测常用的技术是回归分析。

2．5 BI表示/展现技术——信息可视化与Web应用

数据经过分析后得到的即为商业智能。为了使它们直观、简练地呈现在用户面前，需要采用一定的形式表示和发布出来，通常采用的是一些查询和报表工具。不过，目前越来越多的分析结果是以可视化的形式表示出来。这就需要采用信息可视化技术。所谓信息可视化是指以图形、图像、虚拟现实等易为人们所辨识的方式展现原始数据间的复杂关系、潜在信息以及发展趋势，以便我们能够更好地利用所掌握的信息资源。

随着Web应用的普及，商业智能的解决方案能够提供基于Web的应用服务，这样就扩展了商业智能的信息发布范围，享有访问权限的用户只要能够登陆互连网，就可以及时访问自己所关心的内容。作为基于Web的商业智能解决方案，需要一些基本的组成要素，包括基于Web的商业智能服务器、会话管理服务、文件管理服务、调度、分配和通知服务、装载平衡服务和应用服务等。以应用服务为例，需要为用户的查询、分析和报表请求提供浏览器插件。由于商业智能应用涉及到对多个分散在各地的异构数据库的存取，因此还需要CORBA、IIOP协议和Java、XML语言的支持。

**第四篇：BI实验报告**

BI课程

学号：

姓名：王晓郁

专业：计算机科学与技术

日期：

实验报告

0841902112(08中法)2025年5月20日

实验、QL Server BI Step by Step SSIS 3 【开发语言及实现平台或实验环境】

sql server 2025

【实验目的】

（1）利用SSIS把SQL Server中的数据导出与导入。（2）熟悉和掌握实验的各种操作。

（3）能够在实际的事例中灵活运用学到的有关知识。

【实验原理】 1.SQL语句查询 2.SSIS包功能

【实验要求】

（1）利用SSIS实现了数据的导入和导出,把Product表中的数据导出成Excel,然后对产品编号和名称两个字段经过派生的功能进行转换再导入到数据库中。（2）按照实验步骤所示内容认真完成每一个步骤；

【实验步骤】 1.打开Vs.net 2025,选择商业智能项目,然后选择模板中的Integration Service项目,输入项目名称:

创建好后,就在默认的Package包中进行设计。

2.向控制流中添加一个数据流组件(Data Flow Task),双击进入数据流.从左边的工具箱中选择OLEDB数据源(OLE DB Source),可以看到,添加进去后是红色的.双击设置一下连接,如果没有已经创建的链接,则新建一个连接,选定后设置直接访问Produt表。

OK,确认了之后,发现红色的错误提示已经不存在了。

3.接下来直接再添加一个目标数据源,将数据导出成Excel数据格式,所以选择Excel Destination,同样,双击对Excel连接管理器进行配置,配置好文件名称和路径以后(此处选择在首行显示列名,这样会从第二行开始才开始显示数据),如果文件不存在,直接选择下面的新建,创建新的Excel工作表。

4.然后从左边切换到映射,对数据流中的元数据的列和Excel表中的列进行一一映射,因为刚才是自动创建的Excel工作表,所以默认是根据名称对应的.这样我们就完成了对Product产品表的导出,在右边的解决方案中,右击执行包,可以看到绿色执行成功.5.通过设置sql语句导出所有颜色为黑色的产品,让我们把OLEDB数据源的访问模式改成sql命令,然后输入查询语句:

SELECT \* FROM Production.Product WHERE(Color = \'Black\')

6.接下来,我们再将刚才导出的产品数据导入Product表中.再添加Excel Source和OLE DB Destination,其实就是做和导出相反的过程.把 Excel Source的OLEDB连接指向Excel连接管理器。

OLE DB Destination的连接设置成OLE DB Source中的数据库连接,同样,设置成以[表或视图]的方式访问Product表,确定后发现有红色错误。需要将rowguid字段,这两个字段都是数据库中自动生成的.再次确认后会发现已经没有错误,只剩下了黄色的警告,现在暂时不理会这个警告。

7.把刚才生成的Excel文件删除,重新配置Excel连接管理器生成新的空Excel文件(或者把生成的Excel中的数据删除),然后再次运行包,刚才的数据导出仍然正常,数据导入显示的是没有导入任何数据,这是因为在数据中刚才的数据导出和导出并没有先后,所以他们是同步执行的,执行导入时发现里面的数据为空,所以没有导入成功任何数据.不过,尝试着把数据导入的操作直接放在Excel Destination后面是失败的,Excel Destination就是数据流目标,意味着整个流程的结束.(此时Excel Destination中只可定义一个错误输出)。

8.切换到控制流,再添加一个数据流任务,将第一个数据流任何指向这个(鼠标拖拉绿色箭头):

9.双击刚添加的数据导入(已经编辑的数据流组件名称),把刚才的数据流中的组件剪切过来.清空Excel数据再次运行包,有错误 ,看下面的错误信息: An OLE DB record is available.Source: “Microsoft SQL Native Client” Hresult: 0x80040E2F Description: “语句已终止。”.An OLE DB record is available.Source: “Microsoft SQL Native Client” Hresult: 0x80040E2F Description: “不能在具有唯一索引 \'AK\_Product\_ProductNumber\' 的对象 \'Production.Product\' 中插入重复键的行。”.10.暂时通过添加一次转换,在刚才的数据流源和数据流目标中间再添加一个派生列组件(Derived Column ,Updates column values using expressions).添加一个新列NewProductNumber,在Excel中的产品编号后面加1,组成新的产品编号,同样派生出一个新的产品名称Name,因为在数据库中同样也有唯一性约束.11.还要修改OLE DB目标中的映射,将目标列的ProductNumber对应的输入列ProductNumber改成刚才派生的NewProductNumber.将目标列的Name对应的输入列Name改成刚才派生的NewName.清空Excel数据,再次运行包,都变成了绿色,执行全部成功

12．通过对比数据库,确实已经成功的添加进了93行新的数据

【参考文献】

1、SSIS\_StepV2.pdf文档

【心得体会】

1.熟悉了sql server 2025的环境和熟练了SOL语言的使用。

2.通过实验得知，控制流中的数据导出和导入两个组件其实是前后约束条件的,也就是必须数据导出必须成功了才会执行导入(后面会介绍).另外,所谓的数据导出并不会局限于数据库的导入与导出，数据流源和数据流目标都可以是Excel,Flat File(txt,csv),XML,DataReader等连接.也就是说同样可以实现txt导入Excel,或者是XML导入数据库等操作。

**第五篇：汉语拼音bi怎么发音**

汉语拼音中，每一个都有其独特的发音，你知道汉语拼音bi怎么发音吗？不知道也没有关系，我们一起来学习吧。下面是小编帮大家整理的汉语拼音bi怎么发音，希望大家喜欢。

汉语拼音b怎么发音？

在英文中，B(发音为/bi/，bee)是ISO基础拉丁字母表中的第2个字母。

在欧洲和斯堪的纳维亚半岛上，“B”被用来表示“平面”，在计算机科学中，B是信息存储单元——字节的符号。

在工程方面，B是一个级别的单位。

在化学中，B是硼的符号，是一种化学元素。

汉语拼音的第一个声母，读音：波

网络中B是用来骂人的语气词，常见于网络中(网络用语)。常与

2、S等字符组合使用

发音方法

汉语拼音有两种拼读方法：

拼读法：前音轻短后音重。（例：b+a=ba）

直呼法：先做声母的嘴形再做韵母的音。

拼音有声母和韵母。韵母发音时较响亮，声母发音较轻快。

韵母又分单韵母和复韵母，只包含一个元音的，叫单韵母；包含几个元音或者由元音、辅音组合而成的，叫复韵母。汉字音节的韵母，由韵头、韵腹和韵尾3个部分组成。韵头，又称介音；韵腹，称元音；韵尾则有元音和辅音尾两种。以“卡”字而论，q是声母，ia是复韵母。i是韵头；a是韵腹。

本DOCX文档由 www.zciku.com/中词库网 生成，海量范文文档任你选，，为你的工作锦上添花,祝你一臂之力！