

IBM WebSphere Commerce Analyzer



技术参考

版本 5.5

IBM WebSphere Commerce Analyzer



技术参考

版本 5.5

注意

在使用本资料及其支持的产品之前，请阅读第 87 页的附录 B，『声明』中的信息。

第一版，第一修订版（2003 年 9 月）

本版本适用于 IBM WebSphere Commerce Analyzer V5.5 及所有后续发行版和修订版，直到在新版本中另有声明为止。

IBM 欢迎您提出宝贵意见。您可以使用在线 IBM WebSphere Commerce 文档反馈表单发送您的意见，该表单可在以下 URL 获得：

<http://www-3.ibm.com/software/genservers/commerce/rcf.html>

© Copyright International Business Machines Corporation 2000, 2003. All rights reserved.

目录

本书总述	v
本书中使用的约定	v
相关信息	v
支持 Web 站点	vi

第 1 部分 概述 1

第 1 章 什么是 WebSphere Commerce Analyzer? 3

谁使用 WCA?	3
WCA 如何与 WebSphere Commerce 一起工作?	4

第 2 章 业务问题和业务报表 7

检查业务问题的类别	7
---------------------	---

第 2 部分 使用 WCA 服务器 9

第 3 章 维护 WCA 服务器 11

备份 WCA 数据库	11
使用 DB2 控制中心的“删除”菜单除去数据集市	11
确定商店的语言和货币属性	12
更改缺省报表货币	12
除去商店	13
维护参考文本	13
确定 WCA 财政日历的属性	16
进行财政年度修改	19
更改定制 DMS 的缓冲池、表和表空间	19
保持填充产品和概念层次结构表	20

第 4 章 跟踪和记录 21

生成的日志文件	21
TraceLog 查看器	21
跟踪级别	22

第 5 章 提高性能 23

重新组织 WCA 表	23
如果每日新数据非常大, 更新 MAX_SYNCH_MINUTES	23

第 3 部分 WCA ETL 过程 25

第 6 章 WCA ETL 流 27

第 7 章 复制和抽取 29

第 8 章 使用 ETL 驱动程序 33

受支持的命令标记	33
受支持的全局命令	34
受支持的本地命令	34

兼容 DB2 脚本文件的准则	37
INSERT SQL 语句	38
UPDATE SQL 语句	38
组合 SQL 语句的方式	39
在光标方式下运行	39
在 export / import / load 方式下运行	39
错误处理	40

第 4 部分 定制 41

第 9 章 定制 WCA 之前 43

WCA 内部操作	43
扩展和修订	43
对 WCA 的扩展	43
对 WCA 的修订	44
定制方案	44
WebSphere Commerce 实现更改方案	44
数据集市处理方案	45
WCA 数据集中缺少数据的方案	45

第 10 章 定制 WCA 数据集市 49

WCA 数据集市概述	49
数据集市编码标准	49
数据集市定制任务	49
更改表	49

第 11 章 定制抽取 51

抽取编码标准	51
定制 ETL	52
创建定制 .sql 文件	52
定义 SQL 语句	53
添加执行命令	53
添加文件至正确位置	53
创建定制脚本文件	54
添加定制脚本文件错误消息	54
创建用户定义程序以访问 SQL 文件	54
维护表	54
在数据仓库中心控制数据库中添加步骤	55
调度步骤运行	55

第 12 章 定制复制 57

复制概述	57
捕获和应用	57
登台表	58
复制编码标准	58
复制定制任务	58
创建新的复制表	59
增加日志空间	59
计算表空间大小	59
修改复制控制表	61

启动“捕获”过程	62
注册复制过程	62
运行复制步骤	63
调度复制和抽取	63
第 13 章 定制数据挖掘	65
数据挖掘概述	65
用于群集的数据库模型示例	66
数据挖掘定制任务	67
创建数据挖掘模型	67
注册新数据挖掘模型	68
确定最近购买时间、购买频率和购买金额的值	70

附录 A. WCA 参数	71
预定义 WCA 参数的列表	71
创建或更改参数	80
配置后的参数表	81
附录 B. 声明	87
商标	88
词汇表	89
索引	91

本书总述

本书为系统管理员和市场分析员提供关于使用 IBM® WebSphere® Commerce Analyzer V5.5 (又称为 WCA) 的信息。在安装、配置和执行 WCA 服务器的配置后的设置之后, 请使用本书获取关于使用 WCA 的信息, 这些信息所包含的主题如下:

- 如何对 WCA 服务器执行定期维护活动
- 如何提高 WCA 服务器性能
- 如何定制 WCA。

本书中使用的约定

本书使用以下突出显示的约定:

粗体字表示命令或图形用户界面 (GUI) 控件, 例如字段名、按钮或菜单选项。

等宽字表示示例、您输入的文本或屏幕上显示的文本。

斜体字表示新的术语、书籍标题、CD 标签或必须用实际值替换的变量信息。

相关信息

《IBM WebSphere Commerce Analyzer V5.5 技术参考》和《IBM WebSphere Commerce Analyzer V5.5 数据集市参考》可从 WebSphere Commerce Web 站点获取:

- Business Edition:

http://www.ibm.com/software/webservers/commerce/wc_be/lit-tech-general.html

- Professional Edition:

http://www.ibm.com/software/webservers/commerce/wc_pe/lit-tech-general.html

- WebSphere Commerce - Express:

www.ibm.com/software/commerce/express/


《IBM WebSphere Commerce Analyzer V5.5 安装与配置指南》和自述文件也可以从该 Web 站点获取。

注: 《IBM WebSphere Commerce Analyzer V5.5 安装与配置指南》还可以在 IBM WebSphere Commerce Analyzer CD 的 *locale* 目录中获取。在安装后, 有关那些已安装的语言环境的书也可在 WCA 服务器的 %IWDA_DIR%\doc\locale 目录中找到。英文书籍安装在 %IWDA_DIR%\doc\en_US 目录中。

- *locale* 是计算机的语言环境。例如, 对于美国英语, 语言环境是 en_US。
- %IWDA_DIR% 是代表安装有 WCA 的 Windows® 目录的环境变量。在缺省情况下, 该目录是: C:\Program Files\IBM\WCA

下表显示随 WCA 一起提供的文档以及相关描述和文件名。

书	描述	PDF 文件名
《IBM WebSphere Commerce Analyzer V5.5 安装和配置指南》	提供安装和配置 WCA 服务器所需的信息。	install.pdf

书	描述	PDF 文件名
《IBM WebSphere Commerce Analyzer V5.5 技术参考》	<p>提供关于以下主题的信息:</p> <ul style="list-style-type: none"> • 维护 WCA 服务器 • 解决可能发生的问题 • 提高性能 •  使用 IBM DB2® Intelligent Miner™ for Data • 为系统管理员包含有关以下内容的信息: <ul style="list-style-type: none"> – 客户方案 – 扩展数据库模式 (添加到业务报表; 这可以是向 WCA 数据集市添加字段或 (更为常见) 添加报表) – 通过添加到业务报表、添加报表或更改 WCA 提供的模式来定制数据库模式) 	techref.pdf
《IBM WebSphere Commerce Analyzer V5.5 数据集市参考》	包含关于 WCA 数据集市表和视图的信息。	datamrt.pdf
自述文件	提供关于 WCA 的最新信息	README.txt

您可以使用 Adobe Acrobat Reader 查看 PDF 文件。有关 Adobe Acrobat Reader 的信息, 请访问 www.adobe.com。

关于 WebSphere Commerce 的信息, 请参阅以下文档:

- 《IBM WebSphere Commerce 快速入门》
- 《IBM WebSphere Commerce 安装指南》
- 《IBM WebSphere Commerce 基础》

支持 Web 站点

检查 WebSphere Commerce Web 站点以获取 WebSphere Commerce 修订包。任何对 WCA 的更新会包含在 WebSphere Commerce 修订包中:

- www.ibm.com/software/commerce/support/

对本指南的更新可在以下地址找到:

- www.ibm.com/software/commerce/library/

关于 WCA 5.5 中包含的产品的支持信息, 请查看以下 Web 站点:

IBM DB2 通用数据库™企业服务器版

www.ibm.com/software/data/db2/udb/support.html

IBM DB2 Intelligent Miner for Data

www.ibm.com/software/data/iminer/fordata/support.html

第 1 部分 概述

此部分提供了有关 IBM WebSphere Commerce Analyzer V5.5 的概述, 其包含的主题如下:

- 什么是 WebSphere Commerce Analyzer?
- 业务问题和业务报表

第 1 章 什么是 WebSphere Commerce Analyzer?

IBM WebSphere Commerce Analyzer V5.5 (也称为 WCA) 是 WebSphere Commerce 的可选安装的功能。WebSphere Commerce 提供用于创建并维护网上商店的工具。WCA 提供阐述商店中市场营销和购物活动的信息。您可以使用这些信息帮助管理商店的成功运营。

WCA 创建并维护数据集市, 其中包含了生成关于商店的业务报表所需的信息。数据集市是在 WCA 服务器上创建的 IBM DB2 关系数据库。数据集市包含从 WebSphere Commerce 事务数据库服务器抽取并重新组织以便有效报告的数据。IBM DB2 提供数据库管理所需的工具。


业务经理可以从基于浏览器的 WebSphere 贸易加速器 (与 WebSphere Commerce 安装在一起) 访问业务报表。

注: 业务经理可以是市场经理、采购经理或销售经理。

在安装完 WCA 第一次运行复制和抽取过程之后, 从 WebSphere Commerce 数据库到 WCA 数据集市的新数据复制和抽取可以调度为定期运行。例如, 您可以将这些过程调度在午夜之后立即进行, 也可以在正在管理的商店活动较少的其它时候进行。

WCA 可以提供多个商店的信息。WCA 数据集市支持多种语言; 但是, 您所使用的报表应用程序可能限制报表只能使用一种语言。

以下软件随 WCA 一起提供:

- IBM DB2 通用数据库企业服务器版 V8.1
-  IBM DB2 Intelligent Miner for Data V8.1

谁使用 WCA?

有三种类型的 WCA 用户。在本书中, 这些用户称为系统管理员、业务分析员和业务经理。在某些业务中, 可能有两个以上的人要充当这些角色, 并且可能具有不同的称呼。

系统管理员是安装、配置并使 WCA 服务器保持运作的人员。系统管理员还执行以下任务:

- 管理 DB2 和 DB2 仓库中心
- 调度业务报表的生成
- 在 WCA 服务器上执行维护活动, 例如备份
- 诊断并解决可能发生的问题

业务分析员熟知数据分析和数据挖掘。对于 WCA, 下列操作需业务分析员与系统管理员协作完成:

-  使用 Intelligent Miner for Data 的数据挖掘功能
- 定义如何定制业务报表

业务经理从业务角度关心商店的运作。该角色制订市场营销策略并跟踪商店的成功运营。业务经理确定商店所面向客户的类型，并计划活动及其相关的广告。

WCA 如何与 WebSphere Commerce 一起工作？

下图显示了 WebSphere Commerce 和 WCA 的组件之间的关系。

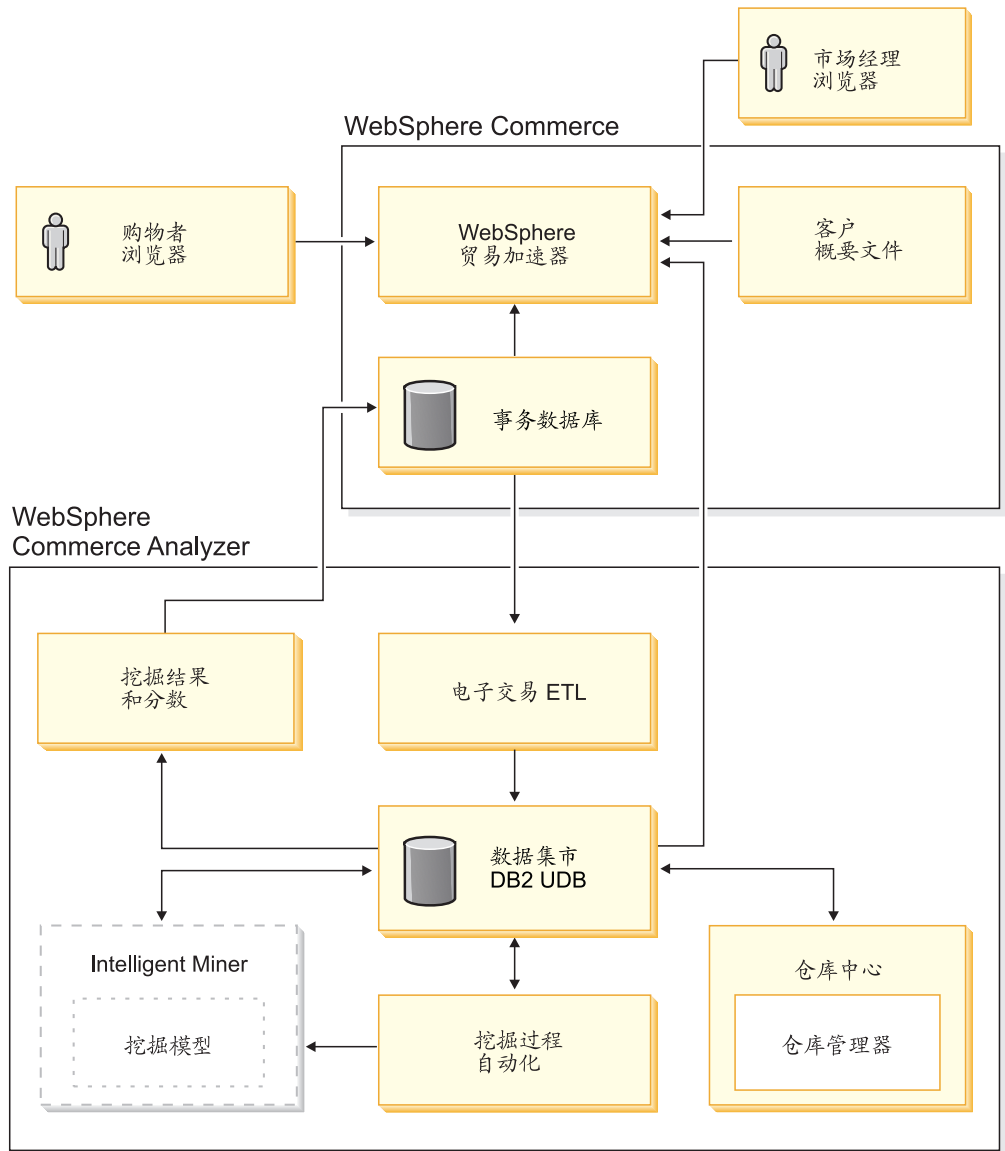


图 1. WebSphere Commerce 和 WCA 组件

购物者使用 Web 浏览器在用 WebSphere Commerce 创建的商店中购物。在购物会话期间，购物者浏览产品，查看显示的广告并可能购买产品。WebSphere Commerce 收集有关查看过的产品和广告以及在每次会话期间购买的商品的详细数据。客户会话数据存储在 WebSphere Commerce 事务数据库服务器上。

WCA 将客户会话数据以及 WebSphere Commerce 事务数据库服务器中的产品和促销数据复制到 WCA 服务器上的临时表中。此操作称为“复制”。WCA 然后将这些数据转

换到可用于报告的表中并将其存储在 WCA 数据集中。此操作称为“抽取”。在初始的复制期间，所有现有的数据都是从 WebSphere Commerce 数据库服务器中获取的。此后，仅获取新的数据。

Professional **Business** IBM DB2 Intelligent Miner for Data 在 WCA 数据集中挖掘数据。它在数据中发现对于回答特定的与数据挖掘相关的业务问题有用的模式，然后将结果存储在 WCA 数据集中。这些结果可从数据集中用于产生与数据挖掘相关的业务报表。

您可以定制业务报表以更好地满足业务需求。有关如何定制业务报表的信息，请参阅《*IBM WebSphere Commerce Analyzer V5.5 技术参考*》。

报表框架是随 WebSphere Commerce 一起提供的报表工具箱，它从 WCA 数据集中的数据产生业务报表。也可以集成其它报表应用程序来产生业务报表。

第 2 章 业务问题和业务报表

业务经理是业务报表的主要用户。本章中的信息将告诉您业务报表可能会阐述哪些类型的业务问题。本章还提供了关于如何查看业务报表的信息，从而帮助确保正确生成报表。

WCA 将数据集市设计基于本书中所记录的业务问题类别之上。由报表应用程序生成的报表可能会针对这些问题的一个子集，并且这些报表还可能针对其它类型的问题。

关于由报表应用程序生成的业务报表的信息，以及关于可能发生的业务报表问题及其解决方案的信息，请参阅随报表应用程序一起提供的文档。

检查业务问题的类别

您可以在以下列表中看到业务报表所阐述的业务问题的类别：

有关 B2B 直销的问题

此类别使您可以确定您的商家到商家交易的效果。此类别中的问题集中于以下几点：

- 销售收入（按帐户概要文件所在的地区）
- 已成为订单的中标报价请求（RFQ）回复的总销售额（按时间段）
- 作为拍卖结果的合同或订单的总销售额
- 占所下订单总数百分比的固定订单数目，以及这些订单相对于整个商店销售额的数额
- 使用各种协商模型的买方的百分比

此类别中的问题例如：根据销售收入、订单频率和退货频率，哪些客户带来的利润多以及哪些客户带来的利润少？

有关产品目录的问题

此类别使您可以分析产品目录条目和出售的产品。此类别中的问题集中于以下几点：

- 按类型和价格出售的产品。
- 未发布的产品目录条目。

此类别中的问题例如：哪些产品目录条目未发布？

有关产品的问题

此类别使您可以检查已放弃产品以及销售额的详细信息。此类别中的问题集中于以下几点：

- 最近出售的 10 个产品。
- 最低的 10 个销售额。
- 关于销售额和所出售产品的详细信息。
- 被人选中然后又放弃的产品。

此类别中的问题例如：销售最少的 10 种产品是哪些？

有关竞销管理的问题

此类别使您可以理解市场营销竞销的性能。此类别中的问题集中于以下几点:

- 在显示的活动中客户所显示的兴趣
- Web 站点竞销的吸引力和效果

此类别中的问题例如: 客户点击过多少次显示的活动?

有关产品顾问的问题

此类别可以评估商店访问者可用的隐喻的有效性。此类别中的问题针对 WebSphere Commerce 产品顾问中的三种隐喻(销售辅助、产品探测和产品比较)。这些问题能够确定每种隐喻的使用情况及其对收入的贡献。

此类别中的问题例如: 使用每种隐喻的购物者的百分比是多少?

有关销售和订单的问题

此类别使您可以了解公司销售交易的趋势, 并评估产生这些趋势的因素。此类别中的问题集中于如下主题:

- 根据价格和售出的数量以及从不同角度对销售结果的查看来评估销售
- 与订单数目和订购总额相关的客户特征
- 订单当前状态

此类别中的问题例如: 已 n 次购买产品 1 的客户的分布情况如何?

有关购物者的问题

此类别可以评估购物者与它们可能下的订单之间的关系。此类别中的问题集中于类似于下面的主题:

- 与订单数量及订购总额相关的购物者特征
- 与所查看的类别及产品相关的购物者特征

此类别中的问题例如: 购物者与平均订购总额之间的关系如何?

有关 Web 站点流量和导航的问题

此类别使您可以分析和确定客户访问商店的趋势。此类别中的问题集中于类似于下面的主题:

- 站点访问者从不同角度的分类
- 将访问者介绍到此站点的人数最多的 Web 站点

此类别中的问题例如: 按照地区将商店访问者分类的结果如何?

第 2 部分 使用 WCA 服务器

此部分讨论如何使用 WCA 服务器。它包含以下主题:

- 维护 WCA 服务器
- 跟踪和记录
- 提高性能

第 3 章 维护 WCA 服务器

以下部分将提供关于在 WCA 服务器上进行维护活动的信息:

备份 WCA 数据库

IBM DB2 提供工具来为归档存储而备份数据库。要获取指导, 请打开“DB2 信息中心”。在“DB2 信息中心”窗口中单击**任务**, 展开**备份和恢复文件夹**, 并选择**备份数据库**。

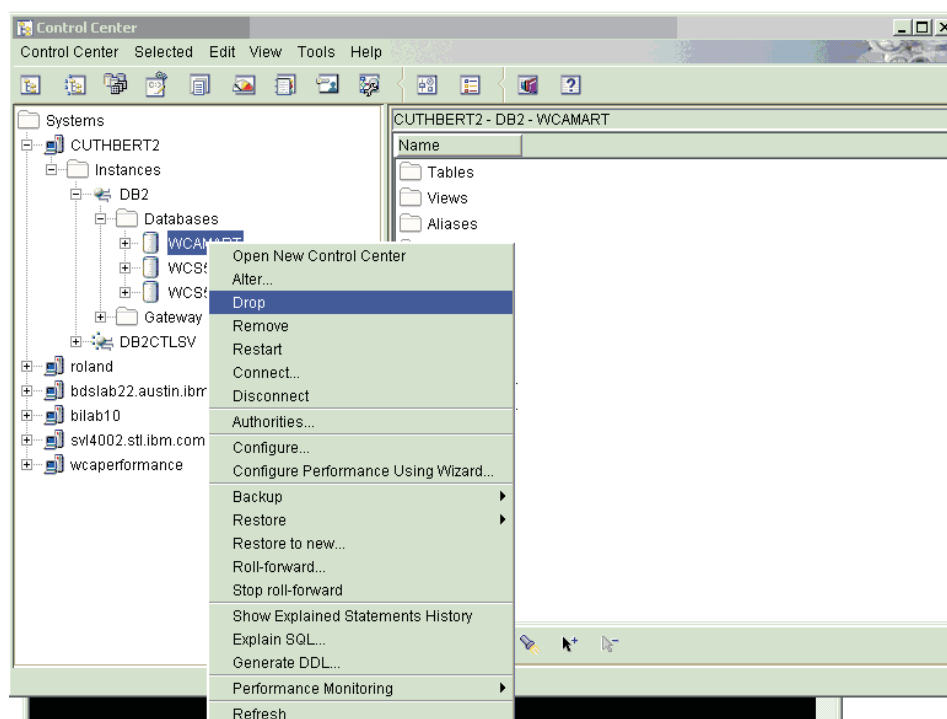
注: 如果使用指导来重新创建数据集市并运行对新数据集市的复制、提升或抽取过程, 则不能使用已备份的数据, 因为同步点已除去。

使用 DB2 控制中心的“删除”菜单除去数据集市

使用**删除**菜单项将数据集市从系统中完全除去。请参阅 DB2 文档以获取关于 DB2 控制中心中**删除**菜单项的更多信息(如下图所示)。

注: 如果您在使用 DB2 控制中心除去 WCA 数据集市, 请不要选择**除去**。**除去**菜单项会除去对数据库的访问权, 但不除去相关联的数据库文件。使用**除去**会导致在使用 WCA 配置工具时出现意外行为。

图 2. DB2 控制中心



确定商店的语言和货币属性

在“选择网上商店以及报表语言和货币”配置步骤中，商店所支持的货币在对应于选中商店的**报表货币**列表框中显示。如果选择了多个商店，则该列表框包含公共的受支持货币。必须提供从每种缺省商店货币到报表货币以及从每种受支持的货币到缺省商店货币的货币转换。

要复查商店支持的语言、货币或货币转换，请查看以下 WebSphere Commerce 表：

CURLIST

标识商店支持的语言。

CURCONVERT

标识商店支持的货币转换。

STORE

标识商店标识。

STOREENT

标识商店的缺省货币。

STORELANG

标识商店支持的语言。

有关商店的货币转换的更多信息，请参阅 WebSphere Commerce 文档。

更改缺省报表货币

如果希望更改商店的缺省报表货币，请执行以下操作：

1. 在用新报表货币更新各种表的时间内，停止运行复制和抽取过程。在将数据更新到新报表货币时，该过程可能在运行复制和抽取过程期间无法完成。
2. 记下 `TIME_CUT_OFF_LOCAL` 参数的参数值。以下命令会更新 `TIME_CUT_OFF_PREV` 参数以便于更改 `REPORT_CURRENCY`：

```
db2 update wca.parameters set (param_value) = '1000-01-01-00.00.00'
```

其中 `param_type=TIME_CUT_OFF_PREV`

3. 启动 WCA 参数管理器。选择**商店**条目以打开“选择网上商店以及报表语言和货币”窗口。从**报表货币**列表框中选择货币并应用您的更改。

注：有关参数管理器的更多信息，请参阅《*IBM WebSphere Commerce Analyzer V5.5 安装和配置指南*》中的『更改配置』

4. 从 IBM DB2 仓库中心运行**有效计算开始**。它将启动以下步骤：
 - 有效计算订购商品
 - 有效计算订单
 - 有效计算事实事件
 - 有效计算事实隐喻
 - 有效计算兴趣列表
 - 有效计算产品定价

除去商店

如果您不再希望产生商店报表，请执行以下过程以除去商店：

1. 确定您不再希望报告的商店的商店标识。
2. 启动 WCA 参数管理器。选择商店条目以打开“选择网上商店以及报表语言和货币”窗口。
3. 清除与商店标识相关联的复选框并应用您的更改。

完成前面的步骤后，复制和抽取过程不再抽取该商店的数据。然而，商店的历史数据仍然存在。如果希望完全除去商店的所有数据，则必须除去基于此商店的事实和尺寸表中的信息。

要从表中除去旧商店数据，请执行以下步骤：

1. 作为数据集市所有者而连接至 WCA 数据集市。
2. 除去事实表中包含希望除去的商店的商店标识的所有行。
3. 除去尺寸表中不直接依赖商店标识的所有行。
4. 除去尺寸表中包含希望除去的商店的商店标识的所有行。
5. 除去 WCA.STORE 表中包含希望除去的商店的商店标识的行。

维护参考文本

当数据集市中的行要求以多种语言进行描述时，WCA 会使用参考表。复制表中的 LANGUAGE_ID 列启用复制工具以限制所显示的语言。WCA.PARAMETERS 表中的 REPORT_LANGUAGE 参数类型与每个参考表中的 LANGUAGE_ID 列结合使用，从而以选定为缺省语言的 WCA 语言来提供信息。WCA ETL 过程也使用此参数从 WebSphere Commerce 映射。

有些参考表从 WebSphere Commerce 数据集市复制，而另一些则使用预定义的属性文件（该文件包含对 WCA 数据集市的附加描述）填充。

在“装入参考和财政阶段文本”配置步骤期间，填充参考表并为装入到阶段表的每个阶段添加阶段描述。

如果已定制 WebSphere Commerce 设置，则在运行 WCA 配置前按要求复查缺省描述并添加或修改参考。

用 %IWDA_DIR%\lib\reftable.properties 文件控制参考表的填充。

特定于语言的文本分别在

%IWDA_DIR%\lib\nls\Reference_Table.properties、%IWDA_DIR%\lib\nls\Reference_Table_locale.properties 文件中定义。

要填充附加的参考表，请修改 reftable.properties 文件并添加适当的插入语句。Reference_Table.properties 文件 %-signs 中所包含的字符串是作为键来使用的，但模式和语言除外，因为它们是自动填充的。请确保继续表键的不间断编号模式。

要对参考表添加附加键，请为所需的所有语言修改 Reference_Table.properties 文件。请确保为 reftable.properties 文件中使用的指定键添加值。例如，如果使用键标识 sdesc 和 ldesc，则您必须为 Reference_Table.properties 文件中的三个键都添加值。

如果缺少参考，则 WCA 抽取过程会在仓库中心的**开始抽取**步骤期间指示添加或删除选中标记失败。

要将缺少的数据包含进来，您可使用“装入参考和财政阶段”窗口以装入修改后的参考，或将它们直接添加至所指示的表。请参阅《*IBM WebSphere Commerce Analyzer V5.5 安装和配置指南*》中的『装入参考和财政阶段』。

下表显示直接映射到 WebSphere Commerce 表中的一组 WCA 参考表。

表 1. WCA 参考表

WCA 表.列名	WebSphere Commerce 表.列名	WebSphere Commerce 的版本
ACCT_STATUS_REF. ACCT_STATUS_ID	ACCOUNT.STATE	<ul style="list-style-type: none"> • WebSphere Commerce Business Edition V5.5
AD_TYPE_REF. AD_TYPE_MAP	CPPMN.PROMODISPTYPE	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
ADJUST_LEVEL_REF. ADJUST_LEVEL_MAP	ORDIADJUST.DISPLAYLEVEL	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
ADJUST_TYPE_REF. ADJUST_TYPE.MAP	ORDIADJUST.CALUSAGE_ID	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
ADDRESS_TYPE_REF. ADDRESS_TYPE	ADDRESS.ADDRESS_TYPE	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
AGE_RANGE_REF. AGE_RANGE	USERDEMO.AGE	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
APV_STATUS_REF. APV_STATUS_ID	MEMBER.STATE	<ul style="list-style-type: none"> • WebSphere Commerce Business Edition V5.5
BUYERPOTYP_REF. BUYERPOTYP_ID	BUYERPO.BUYERPOTYP_ID	<ul style="list-style-type: none"> • WebSphere Commerce Business Edition V5.5
CON_STATUS_REF. CON_STATUS_ID	CONTRACT.STATE	<ul style="list-style-type: none"> • WebSphere Commerce Business Edition V5.5
CPN_STATUS_REF. CPN_STATUS_MAP	CPPMN.STATUS	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5

表 1. WCA 参考表 (续)

WCA 表.列名	WebSphere Commerce 表.列名	WebSphere Commerce 的版本
CPN_TYPE_REF. CPN_TYPE_MAP	CPPMN.PURCHASECONDTYPE	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
DAY_OF_WK_REF	无相应 WebSphere Commerce 表	
DAY_RANGE_REF	无相应 WebSphere Commerce 表	
DR_MEMBER_TYPE_REF	无相应 WebSphere Commerce 表	
EVENT_TYPE_REF	无相应 WebSphere Commerce 表	
GENDER_REF. GENDER	USERDEMO.GENDER	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
GENDER_REF. GENDER	无相应 WebSphere Commerce 表	
HOUR_TYPE_REF	无相应 WebSphere Commerce 表	
INCOME_REF. INCOME_RANGE	USERDEMO.INCOME	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
INVNTY_STAT_REF. INVNTY_STAT_SDESC	ORDERITEMS.INVENTORY STATUS	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
LAST_UPDATED_REF	无相应 WebSphere Commerce 表	
MARITAL_STAT_REF. MARITAL_STATUS	USERDEMO.MARITAL_STATUS	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
MEMBER_TYPE_REF. MEMBER_TYPE	USERS.REGISTERTYPE	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
MPE_TYPE_REF. MPE_TYPE	MPETYPE.NAME	<ul style="list-style-type: none"> • WebSphere Commerce Professional Edition V5.5 • WebSphere Commerce Business Edition V5.5
MPF_TYPE_REF. MPF_TYPE_DESC	USERS.PROFILETYPE	<ul style="list-style-type: none"> • WebSphere Commerce Business Edition V5.5

表 1. WCA 参考表 (续)

WCA 表.列名	WebSphere Commerce 表.列名	WebSphere Commerce 的版本
ORDER_STATUS_REF. ORDER_STATUS	ORDERS.STATUS ORDERITEMS.STATUS	<ul style="list-style-type: none"> WebSphere Commerce Professional Edition V5.5 WebSphere Commerce Business Edition V5.5
ORIGIN_REF. ORIGIN_ID	CONTRACT.ORIGIN	<ul style="list-style-type: none"> WebSphere Commerce Business Edition V5.5
PARTROLE_REF. PARTROLE_ID	PARTICIPNT.PARTROLE_ID	<ul style="list-style-type: none"> WebSphere Commerce Business Edition V5.5
PER_AGGG_REF	无相应 WebSphere Commerce 表	
RANK_RANGE_REF	无相应 WebSphere Commerce 表	
RFQ_ENDRES_REF. RFQ_ENDRES_ID	RFQ.ENDRESULT	<ul style="list-style-type: none"> WebSphere Commerce Business Edition V5.5
RFQ_STATUS_REF. RFQ_STATUS_ID	RFQ.STATE	<ul style="list-style-type: none"> WebSphere Commerce Business Edition V5.5
RSP_STATUS_REF. RSP_STATUS_ID	RFQRSP.STATE	<ul style="list-style-type: none"> WebSphere Commerce Business Edition V5.5
SEG_ATTRIBUTE_REF	无相应 WebSphere Commerce 表	
SEGMENTATION_REF	无相应 WebSphere Commerce 表	
TRADETYPE_REF. TRDTYPE_ID	TRADING.TRDTYPE_ID	<ul style="list-style-type: none"> WebSphere Commerce Business Edition V5.5
TRD_STATUS_REF. TRD_STATUS_ID	TRADING.STATE	<ul style="list-style-type: none"> WebSphere Commerce Business Edition V5.5

确定 WCA 财政日历的属性

在“装入参考和财政阶段文本”配置步骤期间，根据对财政年度开始日期和要装入的阶段数的选择而装入阶段表，并根据语言选择而为每个阶段添加阶段描述。阶段的装入是基于阳历的。

WCA.PERIOD 表包含关于财政年度的每个工作日的信息。WCA.PER_DESC_REF 字段包含有关阶段表中每个阶段标识的特定于语言的描述文本。

注: PER_ID 和 CALENDAR_DATE 字段是 WCA.PERIOD 表中影响抽取过程的仅有字段。不要更改这些字段或列。

以下规则适用于 WCA.PERIOD 表 (缺省属性):

表 2. WCA.PERIOD 表规则

规则	值
WCA 缺省财政日历年度的第一天	1 月 1 日
WCA 缺省财政日历年度的最后一天	12 月 31 日

表 2. WCA.PERIOD 表规则 (续)

规则	值
WCA 的缺省财政日历会注意闰年	2000 年 2 月 29 日、2004 年 2 月 29 日等等。
财政月份的范围与阳历相同	<ul style="list-style-type: none"> • 1 月 1 日 - 1 月 31 日 = 财政年度的第一个月 • 2 月 1 日 - 2 月 28 日或 29 日 = 财政年度的第二个月
财政季度的范围与阳历相同	<ul style="list-style-type: none"> • 1 月、2 月、3 月 = 财政年度的第一个季度 • 4 月、5 月、6 月 = 财政年度的第二个季度
财政月份的工作周、财政季度的工作周和财政年度的工作周的范围。	<ul style="list-style-type: none"> • 2002 年 7 月 31 日: <ul style="list-style-type: none"> - WK_OF_FM = 5 - WK_OF_FQ = 5 - WK_OF_FY = 31 • 2002 年 8 月 1 日: <ul style="list-style-type: none"> - WK_OF_MO = 1 - WK_OF_FQ = 5 - WK_OF_FY = 31

注: 财政年度的工作周、财政季度的工作周和财政月份的工作周的范围对应于阳历 (年度的第一天是 1 月 1 日) 所使用的范围。例如, 如果财政年度 2003 开始于 2003 年 7 月 1 日, 则:

- 2003 年 7 月 1 日: 工作日 = 1、工作周 = 1、月份 = 1、季度 = 1、季度中的工作周 = 1。
- 2003 年 7 月 31 日: 工作日 = 31、工作周 = 5、月份 = 1、季度 = 1、季度中的工作周 = 5。
- 2003 年 8 月 15 日: 工作日 = 46、工作周 = 3、月份 = 2、季度 = 1、季度中的工作周 = 7。

下表显示缺省情况下 WCA.PERIOD 表中的每列是如何填充的:

表 3. WCA.PERIOD 表的缺省填充

列	描述
PER_ID	生成的标识 (从 1 开始)。
PER_AGGR_ID	所有行的缺省值是 1, 它表明这些阶段是这一天的一部分 (例如, 秒、分钟、小时等等)。该列连接至 PER_AGGR_REF 列。
PER_DESC_ID	在缺省情况下, 与所生成的标识相同。它连接至 PER_DESC_REF 列。
CALENDAR_DATE	日期字段, 它对于使用年、月和日的 PER_ID 使用 DB2 日期类型 (例如: 1 = 2000 年 1 月 1 日, 2 = 2000 年 1 月 2 日)。
DAY_OF_WK	该周的工作日, 它每周复位一次。缺省值为 1 = 星期一、2 = 星期二等等。
DAY_OF_WK_ID	该周的工作日标识, 它连接至 DAY_OF_WK_REF 列。缺省值 = DAY_OF_WK。
DAY_OF_FM	财政月份的工作日, 它每月复位一次。
DAY_OF_FM_ID	财政月份的工作日标识, 它连接至 DAY_OF_FM_REF。缺省值是 DAY_OF_FM。

表 3. WCA.PERIOD 表的缺省填充 (续)

列	描述
DAY_OF_FY	财政年度的工作日，它每年复位一次。
DAY_OF_FY	财政年度的工作日标识，它连接至 DAY_OF_FY_REF。缺省值是 DAY_OF_FY。
WK_OF_FM	财政月份的工作周，它每月复位一次。缺省值是：1 = 该月的第一个 7 天段、2 = 该月的第二个 7 天段等等。
WK_OF_FM_ID	财政月份的工作周标识，它连接至 WK_OF_FM_REF。缺省值 = WK_OF_FM。
WK_OF_FQ	财政季度的工作周，它每季度复位一次。缺省值是：1 = 该季度的第一个 7 天段、2 = 该季度的第二个 7 天段等等。
WK_OF_FQ_ID	财政季度的工作周标识，它连接至 WK_OF_FQ_REF。缺省值是 WK_OF_FQ。
WK_OF_FY	财政年度的工作周，它每年复位一次。缺省值是：1 = 该年度的第一个 7 天段、2 = 该年度的第二个 7 天段等等。
WK_OF_FY_ID	财政年度的工作周标识，它连接至 WK_OF_FY_REF。缺省值是 WK_OF_FY。
MON_OF_FY	财政年度的月份，它每年复位一次。缺省值基于阳历月份：1 月、2 月、3 月等等。
MON_OF_FY_ID	财政年度的月份，它连接至 MON_OF_FY_REF。缺省值是 MON_OF_FY。
QTR_OF_FY	财政年度的季度，它每年复位一次。缺省值基于阳历季度：1 月、2 月、3 月 = 1 等等。
QTR_OF_FY_ID	财政年度的季度，它连接至 QTR_OF_FY_REF。缺省值是 QTR_OF_FY。
FISCAL_YR	财政年度。
WEEKDAY_FLG	工作日志。以下是缺省值： 1 星期一 - 星期五 0 星期六 - 星期日
HOLIDAY_FLG	假日标志，它在缺省情况下始终为 0，需要定制。

阶段描述表 WCA.PER_DESC_REF 的填充由 %IWDA_DIR%\lib\refperiod.properties 控制。LOWRANGE 和 HIGHRANGE 定义每个阶段参考表的最低和最高值。

表 4. WCA.PERIOD 表的描述源

列	描述
WK_OF_FM	范围是 01 - 05。
WK_OF_FQ	范围是 01 - 15。
WK_OF_FY	范围是 01 - 55。
MON_OF_FY	范围是 01 - 12。
QTR_OF_FY	范围是 01 - 04。
DAY_OF_FM	范围是 01 - 31。
DAY_OF_FY	范围是 01 - 366。

特定于语言的文本和描述模式分别在 %IWDA_DIR%\lib\nls\Reference_Table.properties、%IWDA_DIR%\lib\nls\Reference_Table_locale.properties 文件中定义。

例如，要填充列 DAYDESC，则使用键 PER_DESC_REF.DAYDESC 来查找该模式。Reference_Table_en_US.properties 文件中的缺省模式是 FY{0}D{4}。花括号中的值用财政年度的实际值（4 位数）和财政工作日的实际值（3 位数）替换，例如 FY2003D035。

您可以更改模式以修改描述，例如，FY{0}-{2}-{9} 将导致 FY2003-02-04（对应于上述示例）。请参阅 Reference_Table.properties 文件以获取可用选项的完整列表。

进行财政年度修改

在 WCA 配置中的“装入参考和财政阶段文本”窗口中，您可以选择财政年度的开始日期和应当装入阶段的年数。

如果您已完成该步骤，但需要更改财政年度的开始日期，则在您第一次运行复制 / 抽取之前执行以下步骤：

1. 关闭 WCA 配置管理器和 WCA 参数管理器。
2. 从 WCA.PERIOD 表和 WCA.PER_DESC_REF 表中删除所有条目。
3. 从 WCA.PARAMETERS 中删除参数类型 FISCAL_YEAR_START 和 FISCAL_PERIODS_UNTIL。
4. 打开 %IWDA_DATA%\tmp 目录中的 StepMgr.prefs 文件。
5. 查找 cfg.fyloaded 和 cfg.fystart 的条目并除去它们。
6. 保存 StepMgr.prefs 文件。
7. 启动 WCA 参数管理员并重新运行该步骤以装入财政阶段。

如果已经运行复制 / 抽取，则必须重新运行 WCA 配置管理器以创建新数据集市（诸如 WCA.FACT_ORDERS 和 WCA.FACT_ORDERITEMS 的几个表中均使用阶段标识）。

更改定制 DMS 的缓冲池、表和表空间

在“创建 WebSphere Commerce Analyzer 数据集市”配置步骤期间，用预定义的缓冲池、表和表空间来设置数据集市。

您可以通过单击**定制**按钮来复查并更改表空间设置。如果您需要对缓冲池和表进行更改，或对表空间设置进行额外的更改，您可以更改用于此步骤的脚本。

以下脚本用于模式 WCA:

```
wca_drp_bufferpools_dms.sql
wca_drp_tbsp_dms.sql
wca_drp_tables.sql
wca_drp_indexes.sql
wca crt_bufferpools_dms.sql
wca crt_tbsp_dms.sql
wca crt_tables_dms.sql
wca crt_indexes_dms.sql
```

以下脚本用于模式 WSA:

```
wsa_drp_bufferpools_dms.sql
wsa_drp_tbsp_dms.sql
wsa_drp_tables.sql
wsa_drp_indexes.sql
wsa_crt_bufferpools_dms.sql
wsa_crt_tbsp_dms.sql
wsa_crt_tables_dms.sql
wsa_crt_indexes_dms.sql
```

脚本的位置取决于 WebSphere Commerce 服务器的版本:

```
%IWDA_DIR%\bin\db2\55be_ext.
```

要修改这些脚本, 请把原始脚本复制到:

```
%IWDA_DATA%\tmp 目录并在运行“创建 WebSphere Commerce Analyzer 数据集市”  
配置步骤之前进行任何必要的更改。
```

保持填充产品和概念层次结构表

如果从 WCA 表删除记录, 不要从本部分的表中除去记录。这些表必须保持完全填充, 以支持产品和概念层次结构表的填充。

支持填充产品层次结构表的表:

- IWH.CATGRP_REL_R
- IWH.CATENTREL_R
- IWH.CATENTRY_R
- IWH.CATTOGRP_R
- IWH.CATGPENREL_R

支持填充概念层次结构表的表:

- IWH.ICKNOWLEDG_R
- IWH.ICKNOWDESC_R

第 4 章 跟踪和记录

WCA 提供高级的跟踪和记录组件（以下称为 *TraceLog*）。TraceLog 主要用于 IBM 支持用途。

TraceLog 带来以下好处：

- 提供对 WCA 可运行组件的全面跟踪和记录
- 检测 WCA 产品错误状况
- 允许 IBM 支持进行远程错误状况诊断
- 将 WCA 日志文件和一些非 WCA 日志文件进行合并

要运行 TraceLog，请在命令提示符下输入以下命令：

```
%IWD_DIR%\bin\runlm.bat
```

注：“启动”菜单也会提供一个称作**日志查看器**的条目，该条目会调用 TraceLog。

生成的日志文件

TraceLog 工具生成三个主要的日志文件：

- **目录日志**

“目录”日志保持对 WCA 运作期间发生的任务、步骤和错误事件的跟踪。

- **配置日志**

可通过在 WCA 配置管理器面板上单击**查看日志**按钮来访问“配置”日志，也可以跳至其所在的目录来访问该日志：

```
~\%IWD_IWDA_DATA%\log
```

- **系统日志**

“系统”日志由 ETL 驱动程序在常规操作期间生成。它包含关于执行了哪些 ETL 步骤且是否产生了任何错误的信息。这些日志以 XML 文档格式存储在单独的每日日志文件中。使用“日志浏览器”程序可查看这些日志。

TraceLog 查看器

WCA 日志管理器使您能查看由 TraceLog 生成的综合的 WCA 日志文件。WCA 日志管理器启动 Web 浏览器，然后使用“TraceLog 查看器”工具展开条目并提供关于在事件发生之时 WCA 系统的状态的详细信息。

“TraceLog 查看器”工具是 Java™ 应用程序，使用导航树结构来组织所有的日志文件。您可以展开或收缩树以缩小和放大关于已记录事件或已记录错误的详细信息，包括查看关键参数的值。目录框架显示按日志类型和日期排序的跟踪日志列表。如果您展开 TraceLog 条目，则将看到任务、步骤和操作。还可看到错误、注释、SQL 命令或文件链接。

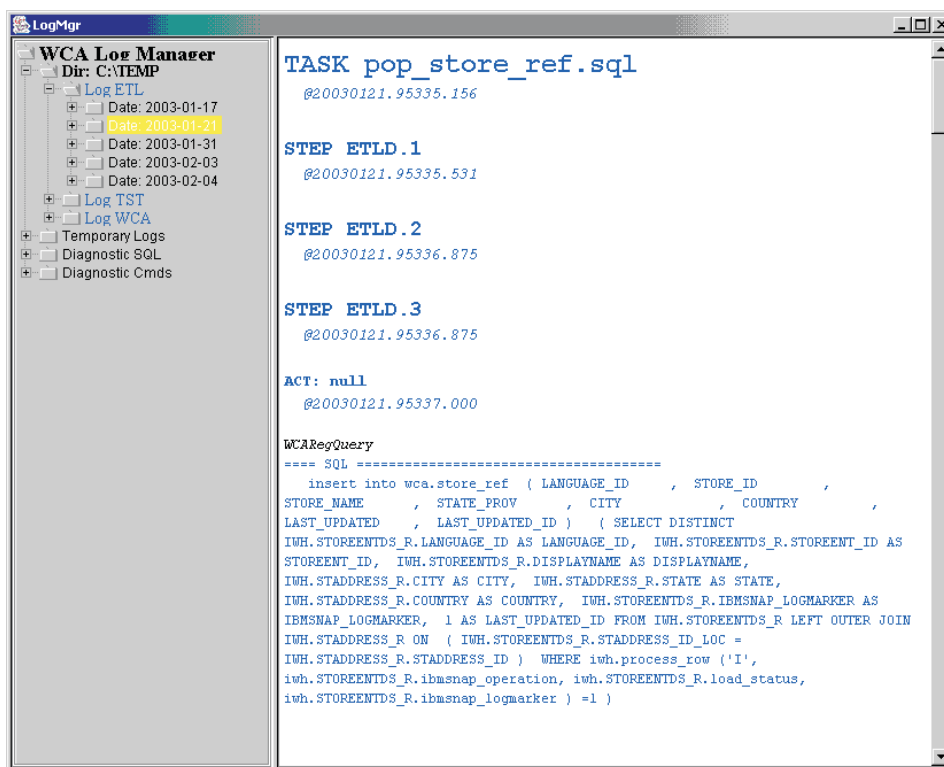


图 3. WCA 日志管理器

跟踪级别

如果需要更多疑难解答信息，您可以通过在以下目录创建一个称为 `tracelog.settings` 的文本文件来打开跟踪：

- `%IWDA_DIR%\log`

如果希望对每个 ETL 步骤打开跟踪，请执行以下操作：

- 用以下条目创建 `tracelog.settings` 文件：`ALLTASKS=1`

如果只想要跟踪某一特定步骤，请执行以下步骤：

- 在 `logmgr` 显示中查看任务标题，然后输入以下行：
 - `pop_store_ref.sql=1`

注：跟踪文件很快就会变得很大。记住：一旦问题解决请关闭跟踪。

第 5 章 提高性能

使用本章中的各节提高 WCA 服务器性能。

重新组织 WCA 表

随着时间的流逝，当对 WCA 数据集中的表作出很多更新时，这些表可能会产生碎片。因此，以后执行对表的操作（例如插入、删除、更新和查询）将会花费更长的时间。如果您注意到查询响应时间增长，则您可以考虑重新组织这些表（特别是在 **RUNSTATS** 命令生成的统计信息是最新的情况下）。

要确定您是否需要重新组织表，请使用 **REORGCHK** 命令。如果此命令的输出指示您需要重新组织表以提高性能，请使用 **REORG** 命令重新组织表。

有关这些命令的详细信息，请参阅 *IBM DB2 Command Reference*。

如果每日新数据非常大，更新 MAX_SYNCH_MINUTES

如果每日新数据的数量非常大（例如超出 1,000,000 条记录），请使用以下信息更新所使用的数据块的大小。如果日志已填满，也可以更新数据块的大小。

如果更改数据的累积量大于数据块的大小，则“应用”程序将把单个的预订周期转换为许多个小周期。同样，“应用”程序将积压数据减少到多个可管理的块。它会重试任何失败的小周期，并缩减数据块的大小以与可用的系统资源相匹配。如果在小周期期间复制失败，“应用”程序将从最后一个成功的小周期开始，重试该预订系列。指定的分钟数确定了数据块的大小。该值存储在 **SUBSCRIPTION SET** 表的 **MAX_SYNCH_MINUTES** 列中。

要更新该值，请在命令提示符下输入以下命令：

```
db2 connect to wca_datamart_name
db2 update asn.ibmnap_subs_set set max_synch_minutes = number_of_minutes
db2 connect reset
```

其中：

WCA_数据集市_name

此参数提供 WCA 数据集市的名称。

number_of_minutes

此参数指定您为分钟数选择的值。从 60 分钟开始，如有需要，请调整该值以提高性能。

例如，您可以输入：

```
db2 connect to wcamart
db2 update asn.ibmnap_subs_set set max_synch_minutes = 60
db2 connect reset
```

第 3 部分 WCA ETL 过程

此部分包含关于 WCA TLT（抽取、转换和装入）过程的信息。ETL 过程就是将数据从某个数据库抽出并将其放入另一个不同类型的数据库时所执行的功能。此部分包含以下信息：

- WCA ETL 流
- 复制和抽取
- 使用 ETL 驱动程序

有关 WCA ETL 的附加信息，请参阅第 52 页的『定制 ETL』和第 44 页的『定制方案』。

第 6 章 WCA ETL 流

您可以确定 WCA 如何运行 IBM DB2 仓库中心中的步骤。 ETL 流包含一个开始步骤和一个结束步骤，并且所有中间步骤均使用 IBM DB2 仓库中心控制的链接而链接在一起。

在 IBM WebSphere Commerce Analyzer V5.5 中，IBM DB2 仓库中心不会告诉您运行这些单独的 ETL 流的顺序。

下表显示此发行版中的每个 ETL 流（按它们运行的顺序）以及每个流的描述。

表 5. ETL 流

执行时间	流序	仓库中心主题名称	仓库中心过程名称	仓库中心步骤名称	描述
手工	-3	性能	ODS runstats	ODS runstats	性能管理流。可以调度或手工运行该流以对数据库执行性能增强操作（例如 runstats）。
对于错误	-2	ETL WebSphere Commerce Analyzer	- 产品层次结构	除去临时表	WCA 产品层次结构错误处理程序流。用于当 X WCA 产品层次结构步骤中发生错误时进行自动清理。
对于错误	-1	Advanced 实用程序	WCA 错误处理程序	WCA 错误处理程序	WCA 错误处理程序流。无论何时在任何 WCA 步骤中发生错误，均会调用该流。它停止仓库服务器并报告错误。
始终	1	Advanced 开始和结束	开始抽取	1. 开始抽取	WCA ETL 开始
始终	2	Advanced 开始和结束	开始复制	2. 开始复制	WCA 复制开始
始终	3	Advanced 开始和结束	开始复制	第一个复制步骤	WCA WebSphere Commerce B2C 复制流
仅对于 B2B	4	Advanced 开始和结束	WC - B2B 复制开始	WC - B2B 复制开始	WCA WebSphere Commerce B2B 复制流
禁用	5	Advanced 开始和结束	WSA 复制开始	WSA 复制开始	IBM Tivoli Web Site Analyzer 复制流

表 5. ETL 流 (续)

始终	6	Advanced 开始和结束	开始抽取		第一个抽取步骤	WCA 标准抽取流
仅对于 B2B	7	ETL WebSphere Commerce Analyzer (B2b)	- B2b - 帐户		X WCSb 帐户	WCA 扩展抽取流
始终	7	ETL WebSphere Commerce Analyzer	- 实际订购商品		X WCSc 报价订购商品	WCA B2C 和 B2B 抽取流，只有当 B2B 禁用时。
始终	8	Advanced 开始和结束	有效计算开始		有效计算开始	WCA 计算流，只有当禁用 IBM Tivoli Web Site Analyzer 支持时。
禁用	8	Advanced 开始和结束	WSA - ETL 开始	WSA - ETL 开始	WSA - ETL 开始	IBM Tivoli Web Site Analyzer 抽取和 WebSphere Commerce 计算流
仅对于 B2B	9	Advanced 开始和结束	WC - B2B ETL 开始	WC - B2B ETL 开始	WC - B2B ETL 开始	WCA B2B 抽取流
仅对于 IM	10	数据挖掘	开始数据挖掘操作	开始数据挖掘操作	开始数据挖掘操作	WCA 数据挖掘培训流
仅对于 IM	11	数据挖掘	开始数据挖掘操作	开始数据挖掘操作	开始数据挖掘应用操作	WCA 数据挖掘应用流
仅对于 IM 和闭环	12	闭环数据挖掘	数据挖掘闭环	数据挖掘闭环	数据挖掘闭环过程	WCA 数据挖掘闭环流
仅对于 IM 和闭环	12	Advanced 开始和结束	结束抽取	结束抽取	结束抽取	如果禁用闭环

第 7 章 复制和抽取

WCA 使用 DB2 replication 将信息从 WebSphere Commerce 数据库表复制到 WCA 服务器上的登台区域中。WebSphere Commerce 是源服务器的一个示例，并且任何驻留在源服务器数据库上的表均称为源表。以下是在使用 DB2 replication 时 WCA 所遵循的一些准则：

- 对应于每个复制的源表，都有一个 WCA 登台表。
- 对应于每个 WCA 登台表，都有一个且仅一个源表。
- 所有 WCA 登台表均以压缩方式复制。
- 对于任何给定的源表，每一行均移入相应的 WCA 登台表。
- 大多数情况下，源表中的所有列均移入相应的 WCA 登台表。
- 每个 WCA 登台表均包含名为 LOAD_STATUS 的列，它指示处理行所用的当前抽取计数。值 -1 表明行尚未得以处理。EXTRACTION_COUNT 是 WCA 参数。所有新行的缺省值是 -1。
- 每个 WCA 登台表均包含名为 IBMSNAP_LOGMARKER 的列。此列指示根据服务器时间而改变行时的时间戳记。除了在第一复制源表期间以外，都是如此。而在第一次复制源表期间，该时间戳记表示复制行所处的时间。此列常与 WCA 参数 TIME_CUT_OFF、TIME_CUT_OFF_PREV、WSA_TIME_CUT_OFF 和 WSA_TIME_CUT_OFF_PREV（取决于源服务器）进行比较。
- 每个 WCA 登台表均包含名为 IBMSNAP_OPERATION 的列。该列的有效值是：I、U 和 D。
 - I** 表示对该行的操作是插入操作。
 - U** 表示该操作是更新操作。
 - D** 表示该操作是删除操作。
- 每个 WCA 登台表均包含名为 IBMSNAP_COMMITSEQ 的列。这是复制处理所必需的，不常为 WCA 使用。
- 每个 WCA 登台表均包含名为 IBMSNAP_INTENTSEQ 的列。这是复制处理所必需的，不常为 WCA 使用。
- 如果源服务器是 WebSphere Commerce，则 WCA 登台表的模式名称是 IWH。
- 如果源服务器是 WebSphere Commerce，则 WCA 登台表名与源服务器表名相同，具有前缀 _R。有些情况下，表名会拼错或由于历史原因而不匹配（例如，源表名在前后发行版之间有更改。）

WCA 抽取的主要目标是获取 WCA 登台表中给出的事务数据库模式，并将它转换为星型模式。WCA 使用一系列松散的准则以完成此任务（根据要求不同，该任务是很灵活的）。大多数的抽取都发生在 DB2 SQL 文件中，这些文件通过 ETL 驱动程序（在下一部分中有记载）运行。事实上，还有些表是通过使用存储过程或外部 Java 代码填充的。

无论是使用 ETL 驱动程序、存储过程或 Java 代码，WCA 抽取在将数据从登台表移入 WCA 数据集市时均使用以下算法。该算法帮助 WCA 确定何时应使用以下因子从登台表处理行。

- 正在执行的查询类型
- 列 IBMSNAP_OPERATION 的值
- 列 IBMSNAP_LOGMARKER 的值
- WCA 参数 EXTRACTION_COUNT (WebSphere Commerce V5.5) 的值
- WCA 参数 TIME_CUT_OFF (WebSphere Commerce 5.5) 的值。TIME_CUT_OFF 是在复制周期开始时设置的。
- WCA 参数 TIME_CUT_OFF_PREV (WebSphere Commerce 5.5) 的值。TIME_CUT_OFF_PREV 是上一次成功抽取的时间。
- WCA 参数 TIME_WINDOW 的值，它是在上一次抽取和上一个复制周期开始之间的时间。

通过调用以下 WCA 用户定义的函数来考虑这些因子:

- 对于 WebSphere Commerce V5.5:

```
IWH.PROCESS_ROW
(OPERATION、IBMSNAP_OPERATION、LOAD_STATUS、IBMSNAP_LOGMARKER)
```

其中:

- OPERATION 是:

- I** 正在寻找抽取时间窗口中尚未插入 WCA 集市的行。
- U** 正在寻找抽取时间窗口中先前已插入 WCA 集市但此后更新过的行。
- D** 正在寻找抽取时间窗口中先前已插入 WCA 集市但最近已删除的行。
- L** 正在寻找抽取时间窗口中的行，无论它们是否已处理。
- N** 正在寻找尚未插入 WCA 集市的行，无论它们是否在抽取时间窗口中。
- A** 正在寻找在当前时间窗口期间尚未插入或更新的行。
- J** 正在寻找上次抽取后尚未插入 WCA 集市的行。
- V** 正在寻找上次抽取后先前已插入 WCA 集市但此后更新过的行。

- IBMSNAP_OPERATION 是:

- I** 特定于复制的变量，表示该行已插入
- U** 特定于复制的变量，表示该行已更新
- D** 特定于复制的变量，表示该行已删除

- LOAD_STATUS 是:

指示任何 WCA ETL 操作上一次处理此行的时间的整数。A -1 表明该行从未处理过。无论何时从登台表处理行，LOAD_STATUS 列都会用当前特定于源的抽取计数参数的值来更新。抽取计数就是在每次 WCA 抽取周期完成时会递增的计数器。对于 Websphere Commerce 5.5 源，该参数是 EXTRACTION_COUNT。

指示任何 WCA ETL 操作上一次处理此行的时间的整数。A -1 表明该行从未处理过。无论何时从登台表处理行，LOAD_STATUS 列都会用当前特定于源的抽取计数参数的值来更新。抽取计数就是在每次 WCA 抽取周期完成时会递增的计数器。对于 Websphere Commerce 5.5 源，该参数是 EXTRACTION_COUNT。


- IBMSNAP_LOGMARKER 是:

IBMSNAP_OPERATION 中指定的操作发生时特定于源的时间戳记。对此的一个例外是在第一次复制表的时候；在这种情况下，它是复制发生期间的特定于源的时间戳记。对于特定源，该列与抽取时间窗口进行比较。

第 8 章 使用 ETL 驱动程序

WCA 在准备 WCA 数据集市的过程中执行许多 ETL 操作。大多数这些操作使用 *ETL 驱动程序* 执行。ETL 驱动程序是这样一个程序：它读取公共 DB2 批处理文件或 SQL 文件，并对正确配置的 WCA 数据集市依次执行每个语句。ETL 驱动程序还通过使用定制 ETL 驱动程序命令标记来增强 DB2 批处理文件，它可以极大提高 SQL 文件的性能，并自动执行一些 WCA 扩展处理要求。ETL 驱动程序可通过 IBM DB2 仓库中心控制数据库中的用户定义程序来访问。所有使用 DB2 批处理文件的 WCA ETL 操作均使用 ETL 驱动程序。

ETL 驱动程序具有以下特征：

- 它用 INSERT、UPDATE、DELETE、DROP、CREATE TABLE、CREATE INDEX、COMMIT、RUNSTATS 和 REORG 命令处理 DB2 批处理文件或 SQL 输入文件。通常，任何不返回结果集的 SQL 语句均是受支持的。除了受支持的 SQL 语句，RUNSTATS 和 REORG DB2 命令也是受支持的。此时没有其它受支持的 DB2 命令。请注意，还可能应用其它限制。
- 它按现状运行 SQL 语句。如果没有为 SQL 语句指定的标记，ETL 驱动程序运行该语句就象从 DB2 命令行运行该语句一样。
- 有了 ETL 驱动程序标记，ETL 驱动程序可以重写查询以执行 CURSOR INSERT、CURSOR UPDATE 和 BULK LOAD 操作。这可以极大提高查询吞吐量。
-  它根据 WCA 挖掘配置面板中的设置来控制 WCA Intelligent Miner 集成步骤的调度。
- 有了正确的 ETL 驱动程序标记，ETL 驱动程序可以自动执行必需的 WCA ETL 处理。例如，它可以更新源表以表明已装入行，或更新目标表上最近更新过的时间戳记。

注：与 WCA 相关联的 SQL 文件中的一些代码行起始如下：

```
--G--
```

或

```
--L--
```

不要修改这些行，这一点非常重要。它们是用于运行 ETL 驱动程序中的 SQL 的特殊命令行。通常不要更改 WCA 中预定义的 SQL 文件。如果更改 SQL 文件，您就是在更改 ETL 函数的逻辑。有关定制准则，请参阅第 52 页的『定制 ETL』。

受支持的命令标记

ETL 驱动程序是通过一组 ETL 驱动程序命令标记而受控的。这些命令标记嵌在 DB2 批处理文件注释中。有两种 ETL 驱动程序命令标记，全局和本地。全局命令标记影响批处理文件中的每个 SQL 语句，并且可在批处理文件中的任一一行中找到。本地命令标记仅影响跟在本地标记后面的第一个语句。

可以使用您自己定制的 .sql 文件中任一 WCA 支持的全局命令或本地命令。

受支持的全局命令

全局命令适用于在指定 .sql 文件中定义的所有 SQL 语句。标识全局命令很容易，因为文件中的该命令行起始于：

```
--G--
```

当前共有两种受支持的全局命令：**CommitStatement** 和 **RunStats**。

CommitStatement

在缺省情况下，ETL 驱动程序不在每个 SQL 语句运行之后发布提交，而是仅在特定批处理文件中的所有 SQL 语句运行之后发布提交。全局命令标记 **CommitStatement** 可用于强制 ETL 驱动程序在每个 SQL 语句运行之后发布提交语句。请注意，有些本地命令强制驱动程序执行提交而不管该设置如何。

此命令的语法如下：

```
<CommitStatement>commit_option</CommitStatement>
```

其中 *commit_option* 具有值 **true** 或 **false**。例如：

```
--G-- <CommitStatement>>true</CommitStatement>
```

RunStats

ETL 驱动程序可将 **Runstats** DB2 命令作为全局命令的一部分运行。但是，建议用户将 **RunStats** 命令作为语句置于 DB2 批处理脚本中，而非使用 ETL 驱动程序的全局命令标记。这允许用户控制准确的 **RunStats** 命令语法以及 **Runstats** 在该文件中运行的时间。

例如，以下命令通知 ETL 驱动程序组成针对 *table_name_r* 目标表的 **RunStats** 命令，然后在运行脚本文件中的其它 SQL 语句之前或之后运行 **RunStats** 命令。

```
--G--<RunStats>  
--G--<TargetTable>table_name_r </TargetTable>  
--G--<RunPosition>position </RunPosition>  
--G--</RunStats>
```

其中 *table_name_r* 表示 **RunStats** 命令的目标表名（例如：iwh.users_r），而其中的 *position* 具有值 **before** 或 **after**。

受支持的本地命令

本地命令仅适用于与其相关联的 SQL 语句。这就是 DB2 批处理文件中在本地标记之后跟随出现的未注释的 SQL 语句。在它适用的语句之前插入本地命令。SQL 语句始终以分号 (;) 定界符字符结束。请注意，如果您决定注释掉标记为本地的 SQL 语句，则建议您也注释掉该语句的本地 ETL 驱动程序标记或将这些标记一起删除。如果您不这样做，则那些标记会应用于下一个未注释的 SQL 语句并会导致意外结果。

您可以使用以下本地命令标记来执行这些类型的操作：

- CursorInsert
- CursorUpdate
- LoadInsert

您可以标识本地命令，因为这些行的前缀是：

```
--L--
```

本地命令标记的模式如下:

```
<ELEMENT CursorInsert (CommitRows,UpdateSource*, LoadOption*, LoadParam*)>
<ELEMENT CursorUpdate (CommitRows, PrimaryKeys, UpdateType, SelectPrimaryKeys*,
UpdateCondition*)>
<ELEMENT LoadInsert (LoadOption, LoadParam*)>
```

请参阅第 39 页的『在光标方式下运行』以获取 `CursorInsert` 和 `CursorUpdate` 操作的示例, 参阅第 39 页的『在 `export/import/load` 方式下运行』以获取 `LoadInsert` 和 `LoadUpdate` 操作的示例。

CommitRows

CommitRows 命令适用于所有使用 **cursor** 的 INSERT 和 UPDATE 情况。

```
<!ELEMENT CursorInsert (CommitRows, UpdateSource*, LoadOption*, LoadParam*)>
```

或:

```
<!ELEMENT CursorUpdate (CommitRows, PrimaryKeys, UpdateType,
                        selectPrimaryKeys*, UpdateCondition*)>
--L-- <CommitRows>num_of_rows</CommitRows>
```

其中 `num_of_rows` 是应在执行增量提交之前处理的行数。

UpdateSource

UpdateSource 标记是用于更新带有 **CursorInsert** 标记的源信息的特殊标记。请参阅关于 WCA ETL 源的部分以获取关于该标记的更多信息。它包含标记 **SourceTable**、**SourceKeys** 和可选标记 **SourceSet** 和 **SourceCondition**。

```
<!ELEMENT CursorInsert (CommitRows, UpdateSource*, LoadOption*, LoadParam*)>
<!ELEMENT UpdateSource (SourceTable, SourceKeys, TargetTable*, TargetKeys*)>
--L-- <UpdateSource></UpdateSource>
```

SourceTable

将 **UpdateSource** 标记中的 **SourceTable** 命令作为 **UpdateSource** 标记的一部分而用于 **CursorInsert** 语句。

```
<!ELEMENT CursorInsert (CommitRows, UpdateSource*, LoadOption*, LoadParam*)>
<!ELEMENT UpdateSource (SourceTable, SourceKeys, TargetTable*, TargetKeys*)>
--L-- <SourceTable>src_table_name</SourceTable>
```

其中 `src_table_name` 是要更新源信息的表的名称。

仅可指定一个源表。

SourceKeys

将 **CursorInsert** 中的 **SourceKeys** 命令作为 **UpdateSource** 标记的一部分而用于 INSERT SQL 语句。

```
<!ELEMENT CursorInsert (CommitRows, UpdateSource*, LoadOption*, LoadParam*)>
<!ELEMENT UpdateSource (SourceTable, SourceKeys, TargetTable*, TargetKeys*)>
--L-- <SourceKeys>srckey1, srckey2, ..... , srckeym</SourceKeys>
```

其中 `srckey1`, `srckey2`, , `srckeym` 是逗号定界列名的列表, 可以唯一标识要更新的源表中的行。

TargetTable

使用 **UpdateSource** 标记中的 **TargetTable** 命令。**TargetTable** 命令是用于指定 ETL 驱动程序的目标表名的方便的方法。表名应使用格式 SCHEMA.TABLENAME。这是一个可选的参数，因为 ETL 驱动程序可从 INSERT 语句自动确定目标表名。

```
<ELEMENT CursorInsert (CommitRows, UpdateSource*, LoadOption*, LoadParam*)>
<ELEMENT UpdateSource (SourceTable, SourceKeys, TargetTable*, TargetKeys*)>
--L--<TargetTable>tgt_table_name</TargetTable>
```

TargetKeys

使用 **UpdateSource** 标记中的 **TargetKeys** 命令。当 **SourceKeys** 命令中指定的列名与目标表中的列名不同时，**TargetKeys** 命令就是必需的。如果使用此命令，则必须有一个在 **TargetKeys** 命令中指定的列对应于 **SourceKeys** 命令中指定的列。

```
<ELEMENT CursorInsert (CommitRows,UpdateSource*,LoadOption*,LoadParam*)>
<ELEMENT UpdateSource (SourceTable,SourceKeys, TargetTable*, TargetKeys*)>
-L--<TargetKeys>tgtkey1,tgtkey2,.....,tgtkeym </TargetKeys>
```

LoadOption

LoadOption 命令共享于 **CursorInsert** 标记和 **LoadInsert** 标记之间。当该命令包含在 **CursorInsert** 标记中时，仅当 ETL 驱动程序确定它可使用 **Export / Load** 选项来替代 **CursorInsert** 时，才会使用该命令。当目标表尚未具有任何行时，会发生这种情况。此后，无论该命令是包含在 **LoadInsert** 还是 **CursorInsert** 标记中，其行为都将相同。

```
<ELEMENT LoadInsert (LoadOption,LoadParam*)>
<ELEMENT CursorInsert (CommitRows,UpdateSource*,LoadOption*,LoadParam*)>
--L--<LoadOption> load_option </LoadOption>
```

其中 *load_option* 仅可为以下值之一：

import

SQL 语句分解为 SELECT SQL 语句和 INSERT SQL 语句。SELECT SQL 语句部分使用 **export** 命令运行，该命令将选中的结果导出至带有定界的 del 格式的外部文件中。

load 根据指定选项，INSERT SQL 语句部分使用 **import** 或 **load** 命令运行。

LoadParam

根据使用 **import** 还是 **load** 命令，确切的参数会有不同。

```
<ELEMENT LoadInsert (LoadOption,LoadParam*)>
--L--<LoadParam>load_parameters </LoadParam>
```

其中 *load_parameters* 是使用 **import** 中的 **modified by** 命令或使用 **load** 命令所设置的参数。

Primary Keys

对 UPDATE SQL 语句使用 **Primary Keys** 命令。如果列出的键与 PrimaryKey 列表中所列的那些键相同，则省略该命令。

```
<!ELEMENT CursorUpdate (CommitRows, PrimaryKeys, UpdateType,
                        SelectPrimaryKeys*, UpdateCondition*)>
<Primary Keys> key1, key2, ....., keym </Primary Keys>
```

其中 *key1, key2,, keym* 是逗号定界列名的列表，可以唯一标识要更新的表中的行。

SelectPrimaryKeys

对 UPDATE SQL 语句使用 **SelectPrimaryKeys** 命令。如果列出的键与 PrimaryKey 列表中所列的那些键相同，则省略该命令。

```
<!ELEMENT CursorUpdate (CommitRows, PrimaryKeys, UpdateType,  
                        SelectPrimaryKeys*, UpdateCondition*)>  
--L-- <SelectPrimaryKeys>key1, key2,...,keym</SelectPrimaryKeys>
```

其中 *key1, key2,.....,keym* 是逗号定界列名的列表，可以唯一标识运行了 SELECT 的表中的行。

UpdateType

UpdateType 命令仅用于 UPDATE SQL 语句。共有两种类型的更新：简单或复合。

```
<!ELEMENT CursorUpdate (CommitRows, PrimaryKeys, UpdateType,  
                        SelectPrimaryKeys*, UpdateCondition*)>  
--L-- <UpdateType>type_of_update_sql</UpdateType>
```

其中 *type_of_update_sql* 仅可为以下类型之一：

简单 表示不用 SELECT 语句而完成对其自身表的更新。例如：

```
update table_name set (column_a,column_b) =  
                    ( 'value_a','value_b')  
    其中 column_c='xxx' 并且 column_d 不为空;
```

复合 表示用一个表中的选择完成对另一个表的更新。例如：

```
update table_name tb set (column_a, ...,column_n) =  
    (select col_a, ....., col_n from table_b_name tb  
     where some_conditions  
     group by some_columns  
    )  
where  
    some_other_conditions
```

UpdateCondition

UpdateCondition 命令仅用于 UPDATE SQL 语句。仅针对特殊条件使用此命令。所列条件附在常规条件之后。

```
<!ELEMENT CursorUpdate (CommitRows, PrimaryKeys, UpdateType,  
                        SelectPrimaryKeys*, UpdateCondition*)>  
--L-- <UpdateCondition>special_upd_conditions</UpdateCondition>
```

其中 *special_upd_conditions* 是必须在 UPDATE SQL 语句分解之后对该语句指定的附加条件。对 PrimaryKeys 中所列键列的常规等同条件会自动添加。

LoadInsert

LoadInsert 标记指定 ETL 驱动程序始终使用“导出/装入”或“导出/导入”选项。首选方法是用 CursorInsert 标记来标记查询，由 ETL 驱动程序确定使用光标插入还是装入方法。然而，仍支持始终执行装入。请参阅在 **CursorInsert** 部分中的 **LoadOption** 和 **LoadParam** 描述。

兼容 DB2 脚本文件的准则

除了公共日志记录机制以外，通过 ETL 驱动程序运行 SQL 语句但不将其标记为光标插入或光标更新，将作用甚微。此部分更详细地描述哪些 SQL 和 DB2 命令受 ETL 驱动程序支持。

INSERT SQL 语句

- 所有已标记的 INSERT 语句必须具有在查询中已清楚定义的目标列。

```
insert into tgt_table (tgt_col1, tgt_col2, ....., tgt_coln)
( select
    src_col1, src_col2, ....., src_coln
    .....
)
```

如果使用了标记, 则

```
insert into tgt_table (select src_col1, src_col2,...,src_coln...)
```

是不受支持的。

- 触发器不会与 DB2 **load** 命令一起工作。如果您使用 **LoadInsert** 操作插入数据, 则任何由触发器填充或修改的列必须在 SQL 语句或在更新语句中手工填充。如有可能, 最好在当前查询中填充这样的列。当驱动程序执行装入时, 它会自动检查以查看目标表是否具有称为 LAST_UPDATED 的 TIMESTAMP 列。如果有, ETL 驱动程序会运行 UPDATE 语句以将此列更新为 CURRENT_TIMESTAMP。此时, ETL 驱动程序将不会自动处理任何其它列。如果使用了以下插入查询:

```
insert into wca.tablename (columnA, columnB)
(select srcA, srcB from iwh.sourcetable)
```

并使用触发器自动填充 columnC, 则 LOAD 操作不会正确填充此列。共有两种填充 columnC 的方法。第一种是用上述 SQL 填充它:

```
insert into wca.tablename (columnA, columnB, columnC)
(select srcA, srcB, srcC from iwh.sourcetable)
```

或执行附加更新查询:

```
update wca.tablename set columnC = (select value from iwh.sourcetable)
```

或

```
update wca.tablename set columnC = (value)
```

UPDATE SQL 语句

- 简单 UPDATE SQL 语句要求将列分组为集合:

```
update tgt_table set (tgt_col1,...,tgt_coln)=(val1,...,valn)
where (...update_conditions...)
```

- 复合 UPDATE SQL 语句要求在 SELECT-UPDATE SQL 语句中定义更新键列:

```
update tgt_table tgt (tgt_col1,...,tgt_coln)
( select
    t1.src_col1,
    t1.src_col2,
    .....
    t1.src_coln
  From
    schema.table_name1 t1
  Where
    tgt.key1=t1.key1 and
    .....
)
```

组合 SQL 语句的方式

组合 SQL 语句可在光标方式下或通过使用 DB2 **export**、**import** 和 **load** 实用程序带增量提交运行。

在光标方式下运行

可在光标方式下带增量提交语句运行 SELECT-INSERT 和 SELECT-UPDATE 类型的组合 SQL 语句。对于 SELECT-INSERT 方式，在用光标完成 INSERT 时会发生对包含键值的源表的附加 UPDATE。

以下是为光标方式类型的操作定义的示例命令集。

CursorInsert

```
<!ELEMENT CursorInsert (CommitRows, UpdateSource*)>
<!ELEMENT UpdateSource (SourceTable, SourceKeys, SourceSet*, SourceCondition*)>
```

带有更新源的光标插入的示例可能如下：

```
--L-- <CurserInsert>
--L--     <CommitRows>1000</CommitRows>
--L--     <UpdateSource>
--L--         <SourceTable>iwh.test_r</SourceTable>
--L--         <SourceKeys>p1</SourceKeys>
--L--     </UpdateSource>
--L-- </CurserInsert>
insert into WCA Advancedet1.test2 (p1,a1,a2,a3)
(select
    p1, a1, a2, a3
from
    iwh.test_r
);
```

CursorUpdate

用光标更新 (CursorUpdate)：

```
<!ELEMENT CursorUpdate (CommitRows, PrimaryKeys, UpdateType,
                        electPrimaryKeys*, UpdateCondition*)>
```

复合光标更新的示例 (其中，选择主键具有与更新主键相同的列名)：

```
--L-- <CurserUpdate>
--L--     <CommitRows>1000</CommitRows>
--L--     <PrimaryKeys>p1</PrimaryKeys>
--L--     <UpdateType>composite</UpdateType>
--L-- </CurserUpdate>
update WCA Advancedet1.test1 t1 set (a1,a2) =
(select
    a1,a2
from
    WCA Advancedet1.test2 t2
where
    t1.p1=t2.p1 and
    t2.p1>2
)
where t1.p1 in (select p1 from WCA Advancedet1.test2 where p1>2);
```

在 export / import / load 方式下运行

使用 DB2 **export**、**import** 和 **load** 实用程序移动大量数据。ETL 驱动程序使您能使用这些 DB2 实用程序来运行组合 SQL 语句 (诸如 SELECT-INSERT 和 SELECT-UPDATE) 并执行增量提交。

SQL 语句分解为 SELECT SQL 语句和 INSERT SQL 语句。SELECT SQL 语句部分使用 **export** 命令运行，该命令将选中的结果导出至外部文件中。根据指定选项，INSERT SQL 部分通过 **import** 或 **load** 命令运行。

以下是为 export / import / load 方式类型的操作定义的示例命令集。

LoadInsert

用 export / load (LoadInsert) 插入:

```
<!ELEMENT LoadInsert (LoadOption, LoadParam*)>
```

装入插入的示例，使用带有以下选项的 **load** 命令:

```
--L-- <LoadInsert>
--L--     <LoadOption>load</LoadOption>
--L--     <LoadParam>modified by forcein</LoadParam>
--L-- </LoadInsert>
insert into WCA Advancedet1.test2 (p1,a1,a2,a3)
(select
  p1, a1, a2, a3
from
  WCA Advancedet1.test3
);
```

LoadUpdate

用 export / load () 更新:

```
<!ELEMENT LoadUpdate (LoadOption, PrimaryKeys, UpdateType,
  SelectPrimaryKeys*, UpdateCondition*, LoadParam*)>
```

复合装入更新的示例:

```
--L-- <LoadUpdate>
--L--     <LoadOption>import</LoadOption>
--L--     <PrimaryKeys>p1</PrimaryKeys>
--L--     <UpdateType>simple</UpdateType>
--L-- </LoadUpdate>
update WCA Advancedet1.test2 set (a2,a3)=(100,'t2a3') where p1<6;
```

错误处理

ETL 驱动程序指定为当发生错误时（在运行 SQL 语句时）或有内部错误时停止。在使用 IBM DB2 仓库中心时，运行 ETL 驱动程序的用户定义程序失败，然后使用 IBM DB2 仓库中心日志记录模式导出错误。当错误发生时，发现问题的最好方法是查看未能找到传递给用户定义程序的实际 SQL 文件名的步骤的定义。如果此文件名是 pop_fact_orderitems.sql，则在 %IWDA_DATA% 目录中有名为 pop_fact_orderitems.log 的日志文件。该文件详细描述 ETL 驱动程序所做的工作，并提供更详细的错误消息。


第 4 部分 定制

此部分包含关于如何定制 IBM WebSphere Commerce Analyzer 的信息。它包含以下主题:

- 定制 WCA 之前
- 定制 WCA 数据集市
- 定制抽取
- 定制复制
- 定制数据挖掘

第 9 章 定制 WCA 之前

WCA 是完全可定制的系统，用户可用新的表和列增强数据集市，向过程添加新的 ETL 步骤或更改关键 WCA 参数。

对定制 WCA 感兴趣的用户在尝试进行任何更改之前，应当具有在 DB2、WebSphere Commerce、replication、DB2 仓库中心、 Intelligent Miner 和 ETL 方面的经验。如果进行了更改，则必须准确记录这些更改，因为它们不会带到 WCA 的较新的版本中。

WCA 内部操作

WCA 包含三个数据库：

- WebSphere Commerce 源数据库，它包含要从中抽取数据以构建 WCA 数据集市的 WebSphere Commerce 客户、产品和订单信息。
- WCA 数据集市目标数据库，它包含所有的客户、产品和订单历史信息。它由 WCA 复制和 ETL 过程持续维护并更新。
- WCA 仓库中心控制数据库，它包含所有的 WCA 和 ETL 步骤定义、WCA ETL 调度和 WCA ETL 流。

WCA 复制过程连续运行，将数据复制至 WebSphere Commerce 和 WCA 数据库上的登台表中。WCA ETL 过程定期将数据从 WCA 登台表抽取、转换并装入 WCA 数据集市。抽取过程还执行附加步骤以更新计算过的字段并为数据挖掘准备数据。

WCA 数据挖掘过程将一组挖掘模型应用于 WCA 数据集市。最终，WCA 闭环步骤将挖掘结果复制回 WebSphere Commerce，以便包含在其概要分析战略中。

扩展和修订

在 WCA 中，共有两种类型的定制：*扩展和修订*。

扩展 该术语指对 WCA 组件的补充（例如，新的报表和新的数据集市表）。此后，使用定制一词时，本手册都是指扩展定义。



修订 此术语指的是对现有 WCA 组件（例如，报表、数据集市和 ETL 过程）的更改。

建议您执行扩展，而不是修订。扩展不会丢失，并会保留在今后对 WCA 的升级中。某些或所有的修订可能会在项目升级时丢失。

对 WCA 的扩展



扩展意味着您可以执行以下操作：

- 更改现有 WCA 生成的报表的副本
- 将新表添加到 WCA 模式
- 将新行添加至 WCA 模式表

- 添加抽取步骤以处理数据抽取
- 添加用户定义的功能以转换已抽取的数据副本
-   更改现有的副本或添加新的 IBM DB2 Intelligent Miner for Data 挖掘库

对 WCA 的修订

修订意味着您可以执行以下操作：

- 调整或更改现有 WCA 生成的报表
- 添加列，或增加 WCA 数据集市表的列大小
- 将新行添加至 WCA 模式表
- 改变 SQL 查询以更好的调整或增强现有的报表
- 更改抽取步骤以抽取和转换 WCA 数据集市表
-   更改现有的 IBM DB2 Intelligent Miner for Data 挖掘库

某些或所有的修订可能会在升级到 WCA 新版本的过程中丢失。

定制方案

如第 43 页的『WCA 内部操作』中所讨论，WCA 的定制可能需要对几个过程点进行更改，以便完成必需定制。下面所列是几个客户方案以及要在各方案中完成的过程点。

在定制 WCA 之前，您必须首先：

- 标识 WCA 数据集市的新要求是什么。
- 对先前记载的 WCA ETL 过程更为熟悉。
- 熟悉仓库中心及其提供的工具。
- 准备记录所有对 WCA ETL 过程所执行的更改。

为了帮助您标识 WCA 数据集市的新要求，请参阅以下方案：

WebSphere Commerce 实现更改方案

方案 1a: WebSphere Commerce 订单处理实现已修改过，不同于所提供的缺省值。这已导致新的订单状态代码集。

- 理解当前的 WebSphere Commerce 实现以及每个新状态代码的含义。特别要理解哪些代码意味着未决订单以及哪些代码意味着应识别收入。
- 请参阅第 13 页的『维护参考文本』以确保 WCA 参考表包含所有具有正确描述的订单状态代码。请遵循此部分中的指导以进行适当的更改。
- 请参阅第 71 页的附录 A，『WCA 参数』以找到处理订单状态代码的所有参数的列表。使用 WCA 参数编辑器进行必需的更改。
- 如果 WCA 已经在处理，请参阅第 12 页的『更改缺省报表货币』。因为在 WCA 中订单状态代码影响几个列的填充，所以您可能希望遵循此过程以根据新状态代码更新那些列。最好从新的 WCA 安装开始并在它上面进行任何更改。

方案 1b: WebSphere Commerce 实现对用户情况调查使用不同于所提供缺省值的分类。

- 理解当前的 WebSphere Commerce 实现以及每个新分类是什么。

- 请参阅第 13 页的『维护参考文本』以确保 WCA 参考表包含所有具有正确描述的状态代码。请遵循此部分中的指导以进行适当的更改。
- 如果 WCA 已经在处理，则根据设计，所有状态代码均应存在，但描述可能不正确。这些描述不影响 WCA 处理，但通常出现在报告中，可能导致使人迷惑的结果。在 WCA 已开始处理之后，始终可以更改参考表中的描述。潜在问题是，同一代码是否会在一段时间后更改其含义。WCA 不处理此例外情况。

数据集市处理方案

方案 2a: 用于确定购物车何时被真正放弃的时间段太短。

- 请参阅第 71 页的附录 A，『WCA 参数』以获取处理在购物车被视为已放弃之前所允许分钟数的所有参数的列表。使用 WCA 参数编辑器进行必需的更改。
- 如果 WCA 已经在处理，请参阅第 12 页的『更改缺省报表货币』。因为已放弃的分钟数影响在 WCA 数据集市填充的标志，所以这些指导对于将当前设置应用于先前记录可能有用。

方案 2b: 用于确定产品的一般报价的方法不足。

- 请参阅第 71 页的附录 A，『WCA 参数』以获取处理产品价格聚集功能的所有参数的列表。使用 WCA 参数编辑器进行必需的更改。
- 如果 WCA 已经在处理，请参阅第 12 页的『更改缺省报表货币』以将这些更改应用到 WCA 数据集市中的当前数据。

方案 2c: 要配属给竞销的订单未被正确配属。

- 请参阅第 71 页的附录 A，『WCA 参数』以获取处理订单状态代码和事实事件表的所有参数的列表。这些参数以 FE 开头。使用 WCA 参数编辑器进行必需的更改。
- 如果 WCA 已经在处理，请参阅第 12 页的『更改缺省报表货币』以将这些更改应用到 WCA 数据集市中的当前数据。

方案 2d: WCA 数据集市中的数据太原始，需要附加处理或转换。

- 理解 WCA 数据集市和 WCA ETL 流。有关更多信息，请参阅第 49 页的『WCA 数据集市概述』和第 27 页的第 6 章，『WCA ETL 流』。
- 标识数据在 WCA 数据集市中的位置。
- 标识在 WCA ETL 处理期间实际装入所提到数据的 WCA ETL 流。
- 标识在其中转换数据的 WCA ETL 流。
- 标识信息应看似如何并如何被使用。
- 标识更改数据的过程，从而使它成为所需数据。
- 标识新数据与哪些事实表（如有）或尺寸表（如有）相关。
- 有关向 WCA ETL 过程添加此步骤的信息，请参阅第 51 页的第 11 章，『定制抽取』。

WCA 数据集中缺少数据的方案

方案 3a: 用户需要 WebSphere Commerce 中和 WCA 数据集中的可用信息，但它未包含在任何报表中。

- 理解 WCA 数据集市和 WCA ETL 流。有关更多信息，请参阅第 49 页的『WCA 数据集市概述』和第 27 页的第 6 章，『WCA ETL 流』。
- 标识数据在 WCA 数据集市中的位置。

- 标识信息应看似如何并如何被使用。
- 请参阅报表工具的文档以确定如何修改报表。

方案 3b: 用户需要 WebSphere Commerce 中的可用信息，但它在 WCA 数据集中不可用。（在此方案中，该信息在 WCA 登台区域中是可用的。）

- 理解 WCA 数据集市和 WCA ETL 流。有关更多信息，请参阅第 49 页的『WCA 数据集市概述』和第 27 页的第 6 章，『WCA ETL 流』。
- 标识数据在 WebSphere Commerce 数据库中所在位置。定位表名。
- 通过使用上述表名来标识数据在 WCA 登台区域中所在位置。定位 IWH 模式中的相应表。
- 标识信息应看似如何并如何被使用。
- 标识新数据与哪些事实表（如有）或尺寸表（如有）相关。
- 有关如何修改 WCA 数据集市的信息，请参阅第 49 页的第 10 章，『定制 WCA 数据集市』。
- 标识应将数据移出登台表的 WCA 抽取流。
- 有关如何向 WCA ETL 过程添加此步骤的信息，请参阅第 51 页的第 11 章，『定制抽取』。
- 请参阅报表工具的文档以确定如何修改报表以包含此新数据。

方案 3c: 用户需要 WebSphere Commerce 中的可用信息，但它在 WCA 数据集中不可用。（在此方案中，该信息在 WCA 登台区域中不可用。）

- 理解 WCA 数据集市和 WCA ETL 流。有关更多信息，请参阅第 49 页的『WCA 数据集市概述』和第 27 页的第 6 章，『WCA ETL 流』。
- 标识数据在 WebSphere Commerce 数据库中所在位置。找到表名。
- 标识应填充 WCA 登台表的复制流。
- 有关如何向新登台表添加此信息的信息，请参阅第 57 页的第 12 章，『定制复制』。
- 标识信息应看似如何并如何被使用。
- 标识新数据与哪些事实表（如有）或尺寸表（如有）相关。
- 有关如何修改 WCA 数据集市的信息，请参阅第 51 页的第 11 章，『定制抽取』。
- 标识应将数据移出登台表的 WCA 抽取流。
- 有关如何向 WCA ETL 过程添加此步骤的信息，请参阅第 51 页的第 11 章，『定制抽取』。
- 请参阅报表工具的文档以确定如何修改报表以包含此新数据。

方案 3d: 用户需要在 WebSphere Commerce 中未提供的信息；它来自外部源。在此场合中，该信息定期更新。

- 理解 WCA 数据集市和 WCA ETL 流。有关更多信息，请参阅第 49 页的『WCA 数据集市概述』和第 27 页的第 6 章，『WCA ETL 流』。
- 标识数据所在位置。
- 标识此信息属于 WCA 登台区域还是直接属于 WCA 数据集市。
- 标识将该数据从源移动至 WCA 登台区域数据集市的最佳方法。（复制、直接 SQL 或其它非数据库移动）。
- 如果使用 WCA 登台区域，则标识应填充此信息的复制流。

- 如果使用 WCA，请参阅第 57 页的第 12 章，『定制复制』以获取关于如何向新登台表添加此信息的信息。
- 标识信息应看似如何并如何被使用。
- 标识新数据与哪些事实表（如有）或尺寸表（如有）相关。
- 有关如何修改 WCA 数据集市的信息，请参阅第 49 页的第 10 章，『定制 WCA 数据集市』。
- 标识应将数据从登台表或其它数据源移动至 WCA 数据集市的 WCA 抽取流。
- 有关如何向 WCA ETL 过程添加此步骤的信息，请参阅第 51 页的第 11 章，『定制抽取』。
- 请参阅报表工具的文档以确定如何修改报表以包含此新数据。

方案 3e: 用户需要 WebSphere Commerce 中未提供的信息；它来自外部源。在此场合中，该信息仅以不太频繁的频率更新。

- 标识数据所在位置。
- 标识将该数据从源移动至 WCA 登台区域数据集市的最佳方法。（复制、直接 SQL 或其它非数据库移动）。
- 标识信息应看似如何并如何被使用。
- 标识新数据与哪些事实表（如有）或尺寸表（如有）相关。
- 有关如何修改 WCA 数据集市的信息，请参阅第 49 页的第 10 章，『定制 WCA 数据集市』。
- 请参阅报表工具的文档以确定如何修改报表以包含此新数据。

方案 3f: 用户希望将新挖掘模型合并至 WCA 实现中。

- 有关如何定制数据挖掘的信息，请参阅第 65 页的第 13 章，『定制数据挖掘』。

第 10 章 定制 WCA 数据集市

此部分讨论如何定制 WCA 数据集市。将讨论的主题如下：

- 数据集市编码标准
- 数据集市定制任务
- 记录更改

WCA 数据集市概述

WCA 将来自 WebSphere Commerce 登台表的运作数据转换为多维格式，并将其存储在 WCA 数据集中。

WCA 数据集市已发布并可定制。请参阅《*IBM WebSphere Commerce Analyzer V5.5 数据集市参考*》以获取有关以下内容的完整描述：

- WCA 表和视图

数据集市编码标准

任何时候均可以定制数据集市。必须遵循几条规则和准则以保持数据集市的运作完整性。

在定制数据集市时请记住下表：

- 不要更改 WCA 表的主键定义。
- 不要在 WCA 表中添加或删除约束。
- 不要删除任何 WCA 表、视图、列或行。
- 不要改变任何按子句连接、分类或分组的 WCA 视图。
- 不要更改 WCA 表中的列的含义：
 - 不要更改数据类型。
 - 不要更改其数据的设置，也不要任何设置将其覆盖。
- 不要删除 WCA 表中的任何索引，因为这可能会影响程序效率。
- 不要除去任何 WCA 用户定义的功能（UDF）。
- 不要除去任何 WCA 存储过程。
- 不要除去 WCA 缓冲池或表空间。

数据集市定制任务

要使多个数据集市更改自动进行，您应创建特定于平台的脚本编制文件以使这些类型的更改自动进行。

更改表

此部分包含在对 WCA 数据集中的表进行更改之前您必须知道的信息。

添加表

当向 WCA 数据集市添加表时，请使用前缀 *UX* 将表名标识为 WCA 数据集市的用户扩展。例如，当添加带有附加成员信息的表时，请输入：

```
create table UX_MEMBER (  
    member_id bigint not null,  
    custcode bigint not null,  
    primary key (member_id)  
)
```

删除表

不要删除任何 WCA 表、视图、列或行。

向现有表添加列

当向现有 WCA 表添加列时，请使用前缀 *UX* 将列名标识为 WCA 表的用户扩展名。

运行以下 SQL 语句以将用户定义的 *UX_MYMEMBER_ID* 列添加至 *PRODUCT* 表：

```
alter table PRODUCT add UX_MYMEMBER_ID big;
```

例如，如果您希望将列 *MEMBER_ID* 添加至 *PRODUCT* 表，您将输入：

```
alter table MEMBER add column UX_CUSTCODE bigint not null;
```

在很多情况下，添加列需要对 WCA 复制和 WCA ETL 过程进行附加更改，使列得以填充。

更改现有表中的列

不要更改 WCA 表中的列的含义：

- 不要更改数据类型。
- 不要更改其数据的设置，也不要任何设置将其覆盖。

记录更改

如果修改或附加数据，则必须记录更改和这些更改的所有权。在 *LAST_UPDATED_REF* 表中记录标识。请确保选择唯一的整数标识（*LAST_UPDATED_ID*）和唯一的描述标识。

第 11 章 定制抽取

如果希望定制 WCA ETL 过程，您必须首先考虑以下情况：

- 决定想要在 WCA 数据集中放置哪些数据。
- 决定必须创建哪些表和 / 或哪些列包含这些数据。
- 确定数据源。
- 遵循 WCA 约定和限制以开发 SQL 来执行任务。
- 创建新的 WCA 过程步骤。
- 将过程步骤链接至适当的 WCA ETL 过程流。

决定需要哪些数据是用户的责任。通常，数据来自以下三个源之一：

- WebSphere Commerce 事务数据库
- WCA 复制登台表
- 一些其它的外部源

在 WebSphere Commerce 事务数据尚未复制到 WCA 登台表中的情况下，请参阅第 57 页的第 12 章，『定制复制』以获取关于如何设置复制所必需的表、表空间和注册的信息。

在先前所列的所有情况下，您必须开发可将数据抽取、转换、移动并装入至 WCA 数据集市的 SQL。这可能需要创建附加的临时表、视图或文件以帮助转换和传送。遵循 WCA 数据集市约定，以防中断任何现有 WCA ETL 过程，并使定位定制变得轻松。

现在必须创建 WCA ETL 过程步骤，并确定此步骤应为哪个 ETL 过程流的一部分。您的步骤应始终链接在所选流的倒数第二步与最后一个步骤之间。这确保定制步骤的所需数据已经处理。请记住每个流的最后一个步骤必须保持为该流的最后一个步骤，这也很重要。不要在任何流的最后一个步骤之后链接定制步骤。当前不能创建新的流。这确保与执行您的步骤相关的信息存在且已就绪，可供使用。

以下是一个常规示例：如果您的步骤需要 WCA 数据以完成数据转换，且 WCA 数据集市是一些效果计算步骤的结果，则您的 ETL 步骤应链接至跟随在效果计算过程流之后的 ETL 过程。

WCA ETL 驱动程序实用程序将作为在 DB2 仓库步骤中使用 SQL 的一种备用方法而提供，用以完成数据 ETL。WCA ETL 驱动程序提供许多有用的功能，包括创建无需附加定制就可运行的所有步骤的详细日常跟踪日志。有关如何构建 ETL 驱动程序步骤的信息，请参阅第 33 页的第 8 章，『使用 ETL 驱动程序』。

抽取编码标准

在使用抽取过程时必须遵循以下编码标准：

- 当您可以向 WCA 数据集市添加自定义功能时，请为客户定义的功能名称加前缀 *UX* 以将其标识为 WCA 数据集市的用户扩展。
- 当您可以向 WCA 数据集市添加存储过程时，请为存储过程名称加前缀 *UX* 以将其标识为 WCA 数据集市的用户扩展。

- 创建特定于平台的脚本文件以使这些类型的更改自动进行。在 Microsoft® Windows 平台上，这些文件被称为 *.bat*（批处理）文件。
 请为每个脚本名称加前缀 *UX*。在命名任何环境变量、临时管道、文件或目录时，请在其名称中使用前缀 *UX*。
 请为脚本文件错误消息加前缀 *UX*，以帮助诊断 WCA 问题。请使用错误消息方案：**UX-nnnnY**，其中 **UX** 是 WCA 错误消息的前缀，*nnnn* 是 4 位错误码，*Y* 将用 **F**（致命错误）、**E**（错误）、**W**（警告）、**I**（信息）替换。
- 创建 IBM DB2 仓库中心脚本以使客户定义的扩展功能的数据抽取、转换、移动和装入（ETL）过程自动进行。为客户定义的步骤、功能和外部调用的脚本文件名加前缀 *UX*。
 可以将步骤添加到 WCA ETL 过程的结束部分，以使它们在相同的时间周期内运行。不要同时运行它们，否则会竞夺系统资源。
 可以添加独立于 WCA ETL 过程的步骤。这些脚本一定不能与 WCA ETL 脚本同时运行，因为它们会影响产品性能。

定制 ETL

本节讨论如何定制 ETL。包含以下主题：

- 创建定制 *.sql* 文件
- 定义 SQL 语句
- 添加执行命令
- 添加定制 *.sql* 文件至正确位置
- 创建定制脚本文件
- 添加定制脚本文件错误消息
- 创建用户定义程序以访问 SQL 文件
- 维护表
- 在数据仓库中心控制数据库中添加步骤
- 调度步骤运行

创建定制 *.sql* 文件

与其创建单个 SQL 语句以将数据从源转换至目标，从而更新数据集市表，您不如创建包含多个作为 SQL 语句的 DB2 更新命令的 *.sql* 文件。

在这个示例中，DB2 命令更新 PRODUCT 表中的 UX_MEMBERID 列：

```
UPDATE WCA.PRODUCT SET UX_MEMBERID=(SELECT UX_MEMBERID FROM WCA.PRODUCT_TMP
WHERE WCA.PRODUCT.PRODUCT_ID= WCA.PRODUCT_TMP.PRODUCT_ID AND
WCA.ORDERITEMS_TMP.LASTUPDATE>=TIMESTAMP(WCA.PARAM_VALUE('TIME_CUT_OFF_PREV'))
WHERE PRODUCT_ID IN (SELECT PT.PRODUCT_ID FROM WCA.PRODUCT_TMP PT
WHERE PT.LASTUPDATE >= TIMESTAMP(WCA.PARAM_VALUE('TIME_CUT_OFF_PREV')))
```

或者运行以下 DB2 创建命令在 DB2 数据仓库中心控制数据库中创建临时表，并从联机数据库中抽取数据。在以下示例中，创建并随后填充用户定义的表 UX_ACME_CUST_MAP 表。

```
CREATE UX_ACME_CUST_MAP TABLE
(
  UX_ACMECARD_ID INTEGER NOT NULL,
```

```
UX_CUSTTYPE INTEGER NOT NULL,  
UX_GEOCODE INTEGER NOT NULL,  
PRIMARY KEY (UX_ACMECARD_ID)  
)
```

定义 SQL 语句

要定义并添加命令集至 .sql 文件中的 SQL 语句，请执行以下步骤：

1. 在纯文本文件中输入希望运行的 SQL 语句。
2. 用分号 (;) 定界符结束每个 SQL 语句。
任何 SQL 语句（不带返回结果集）均可包含在文件中。而且，DB2 命令（如 RUNSTATS 和 REORG）是有效条目。
3. 用文件扩展名 .sql 保存文件。
4. 通过输入以下内容从而在 DB2 命令行窗口中测试 .sql 文件：

```
db2 -tvf sql_filename.sql
```

添加执行命令

要添加一个或多个执行命令，请执行以下步骤：

1. 仅对于 *SELECT-INSERT* 和 *UPDATE SQL* 语句：如果 SQL 语句包含大量数据事务，则必须为 SQL 语句添加附加执行命令标记。否则，跳至步骤 3。
2. 在添加附加命令后，退出所保存的 .sql 文件。
3. 在 SQL 语句中检查使用该功能的任何特殊要求。
4. 对 .sql 文件中的全局和本地命令使用特殊命令符号。
 - 通过以下标记开头的行标识全局命令：

```
--G--
```


如有需要，添加全局命令标记。可提供多个命令行。
 - 通过以下标记开头的行标识本地命令：

```
--L--
```


对所需的 SQL 语句添加本地命令标记。本地命令仅适用于与其相关联的 SQL 语句。在它适用的 SQL 语句之前插入本地命令。
5. 确保用分号 (;) 定界符结束每个 SQL 语句。
6. 在编辑后保存 .sql 文件。
7. 使用命令行环境测试 .sql 文件。

添加文件至正确位置

要将 .sql 文件添加至正确位置：

1. 在生产期间，不同数据源的不同版本文件保存在不同子目录下。当前目录是这样的一个目录，其中：

```
IWDA_DIR
```


标识装有 WCA 的目录。在安装期间设置该目录位置。

```
wcs_source
```


标识 WebSphere Commerce 数据源所在目录。
2. 根据 WebSphere Commerce 数据源的版本，将 .sql 文件保存在正确目录下。

创建定制脚本文件

可以创建定制的特定于平台的脚本文件，您使用该文件使更改自动进行。例如：

1. 通过输入以下 SQL 语句并将 WCA 数据库名称 (WCA1001) 和 WCA 数据集市的用户名和密码指定为参数，从而连接至 WCA 数据集市。

```
CONNECT TO WCA1001 USER USERNAME USING PASSWORD
```

2. 运行 DB2 **-tvf** 命令以创建并填充用户定义的表：

```
db2 -tvf newupdates.sql
```

3. 断开与数据库的连接。

在 Windows 平台上，这些脚本文件被称为 *.bat* (批处理) 文件。

请为每个脚本名称加前缀 UX。在命名任何环境变量、临时管道、文件或目录时，请在其名称中使用前缀 UX。

添加定制脚本文件错误消息

请为脚本文件错误消息加前缀 UX，以帮助诊断 WCA 问题。请使用错误消息方案：**UX-nnnnY**，其中 **UX** 是 WCA 错误消息的前缀，**nnnn** 是 4 位错误码，**Y** 将用 **F** (致命错误)、**E** (错误)、**W** (警告)、**I** (信息) 替换。

可通过向数据仓库中心反馈文件写消息来处理错误，该文件由环境变量 \$VWP_LOG 标识。该文件具有以下格式：

```
<RC>Integer</RC>  
<MSG>Text</MSG>  
<COMMENT>Text</COMMENT>
```

在开始时应除去该文件，仅当发生错误时创建它。如果整数为非零整数，仓库中心会假定发生错误。

创建用户定义程序以访问 SQL 文件

DB2 数据仓库中心中的所有 ETL 步骤均作为对用户定义程序的 ETL 驱动程序的调用而实现。可以创建 IBM DB2 仓库中心脚本以使客户定义扩展功能的数据抽取、转换、移动和装入 (ETL) 自动进行。为客户定义的步骤、函数和外部调用的脚本文件名加前缀 UX。

要创建 IBM DB2 仓库中心用户定义程序，须提供步骤名称作为参数。

维护表

要在 IBM DB2 仓库中心中创建步骤以维护表，您有两个选项：

- 使用 IBM DB2 仓库中心调度该过程

可以通过使用 IBM DB2 仓库中心定期调度该过程来添加独立于 WCA ETL 过程的步骤。然而，IBM DB2 仓库中心必须在定期的数据集市抽取过程运行之前完成该步骤。这些脚本一定不能与 WCA ETL 脚本同时运行，因为它们会影响产品性能。

如果这两个过程同时运行，则性能将受到负面影响。

- 将该步骤附加至标准抽取过程

可以将此步骤附加到标准抽取过程中，使它在 WCA ETL 过程结束时运行。不要同时运行它们，否则会竞夺系统资源。

在作出决定前，复查“业务智能教程”。单击**开始** → **程序** → **IBM DB2 安装工具** → **第一步** →。在“第一步”窗口，单击**使用教程**并单击**业务智能教程：介绍数据仓库中心**。复查以下课程：

- 定义数据转换和移动
- 测试 IBM DB2 仓库中心步骤
- 调度 IBM DB2 仓库中心过程

在数据仓库中心控制数据库中添加步骤

要在数据仓库中心控制数据库中添加步骤，请执行以下步骤：

1. 登录数据仓库中心控制数据库。
2. 在“过程”窗口中，打开现有过程，或创建新过程。
3. 在显示的列表中单击外部程序图标。
4. 选择 **WCA 抽取** → **WCAUPD**，然后添加在“过程”窗口中选定的外部程序。
5. 添加在**仓库目标** → **Advanced 目标表** → 表列表中选定的目标表。
6. 在所添加的外部程序到所添加的目标表之间添加数据链接。
7. 双击新添加的外部程序图标。对应于该新外部程序的“属性”窗口打开。
8. 在“用户定义程序”窗口中，在**名称**字段中提供有效名称。
9. 在“参数”窗口中，编辑以下两个参数的值：
 - 对于 SQL 参数，添加 .sql 文件的名称。
 - 对于 Option 参数，输入值 **Y**、**N** 或一些其它字符串。如果您提供值 **Y**，则 ETL 驱动程序在文件中的每个 SQL 语句运行之后自动提交事务。否则，运行 .sql 文件中的所有 SQL 语句而不进行任何特别提交，除非在需要 **commit** 语句的时候插入 **commit** 语句。
10. 在“过程选项”窗口中，选择正确的代理程序站点。使用。
11. 单击**确定**以保存更改后的属性信息。
12. 在“过程”窗口中，保存所添加的外部程序步骤，然后关闭窗口。

调度步骤运行

要调度步骤运行，请执行以下步骤：

1. 按执行序列在正确的位置添加步骤。
2. 将该步骤提升至“测试”方式，并为该步骤运行测试。
3. 如果测试成功，则将该步骤提升至“生产”方式。在提升成功后，该步骤已就绪，可用于已调度的执行。

第 12 章 定制复制

WCA 为以下类型的电子交易区域处理数据复制和抽取：帐户和合同、竞销和活动、日历（时间段数据）、产品目录层次结构、赠券和折扣、成员、产品顾问隐喻、报价、订购商品和订单以及报价请求。可以从 WebSphere Commerce 以外的源收集电子商业事务数据，如 Web 站点点击流数据。

IBM DB2 仓库中心执行数据合并和集成。可将所有数据源收集并集成到 WCA 数据集中。WCA 使用数据挖掘技术提供对数据的分析。最终，WCA 数据集市成为报表的主数据源。报表应用程序允许您创建并查看合并数据挖掘结果的报表。

复制概述

WCA 使用 DB2 复制技术将 WebSphere Commerce 事务数据复制到 WCA 数据库。WCA 将从 WebSphere Commerce 复制而来的数据转换为多维格式，然后将复制数据存储到 WCA 数据集中。因为事务数据可能会非常多，并且分析处理也可能很密集，因此出于性能原因，WCA 应位于与 WebSphere Commerce 不同的机器上。

WCA 数据库不仅允许 WCA 数据集市模式，而且将登台区域用于复制和任何辅助数据集市。辅助数据集市由可能由报表应用程序引进的表和视图组成。

捕获和应用

DB2 复制技术有两个主要步骤：

捕获 “捕获”程序将源表上发生的所有更改（插入、更新和删除）收集在存储在 WebSphere Commerce 数据库中的更改数据表中。如果“捕获”不在运行，则不会在复制源系统上捕获到对注册源表的更改。

更改数据表存储在表空间中。当表空间到达大小限制时，“捕获”将无法捕获任何附加更改。每个更改数据表有其自身的表空间。

应用 “应用”程序将更改数据带到目标数据库中并将数据存储在登台区域中。对于 WCA，“应用”程序位于目标数据库系统（WCA 数据集市）上。“应用”程序的控制表可以位于本地系统，也可以位于远程系统。当“应用”正在目标数据库所在的同一机器上运行时，它会从源中抽出数据。当“应用”正在不同于目标数据库所在的远程机器上运行时，它将数据推到目标中。“应用”一次仅能传播有限的几行。“应用”会自动将需要传播的对象在几个小周期内分割为几个小预订。

通常，“捕获”和“应用”可以是按需（调度为在指定时间发生），也可以是连续（在一段时间内平均分配工作）。WCA 使用连续“捕获”。在复制周期结束后，另一周期立即紧随，而不是等待连续“应用”。按需“捕获”和“应用”帮助在一段时间内分配复制负载。

登台表

登台表是 WebSphere Commerce 数据源的副本。使用 IBM DB2 仓库中心的复制工具设计复制步骤，从而将数据填充登台表中。WCA 控制数据库主管所有复制和抽取步骤，并协调它们的调度和执行。

WCA 将大量 WebSphere Commerce 表复制到其登台区域中。复制数据存储在 WCA 机器上的登台表中。然后将复制数据作为抽取、转换、移动和装入 (ETL) 的源。登台表使用 ETL 步骤将数据填充 WCA 数据集市。

所有登台表均基于压缩复制选项。已压缩的登台表仅包含对应于源表每行的最近的值。未压缩的表包含对源表每行所作的所有更改，表示每一行的更改历史记录。

临时登台表的名称与源表名称相同，或与源表名称紧密关联。例如，ADDRESS 表的临时登台表是 ADDRESS_R。

每个登台表均有 **LOAD_STATUS** 字段。装入状态检查数据是否已插入 WCA 登台表中。该字段是 WCA 系统定义的在添加数据时使用的字段。

另外，IBM DB2 仓库中心复制工具将这些字段自动添加至每个登台表中：

IBMSNAP_INTENTSEQ

唯一标识更改的日志或日志记录序列号。该值全局升序。

IBMSNAP_OPERATION

字符值 I、U 或 D 分别表示插入、更新或删除记录。在 ETL 步骤处理期间，此字段中任何具有 **time_cut_off** 和 **time_cut_off_previous** 之间时间信息的数据均表明数据插入或更新。

IBMSNAP_LOGMARKER

源服务器上的大概提交时间。此列在完全刷新后始终为空。增量数据处理基于此字段的时间信息。

IBMSNAP_COMMITSEQ

所捕获的 COMMIT 语句的日志记录序列号。

复制编码标准

对于 WCA，所有复制步骤必须包含前缀 **R WCS_**、**R WSA_**、**UX_**。

因为复制和抽取过程在执行数据抽取和转换时是互相配合的，还可参阅第 51 页的『抽取编码标准』。

复制定制任务

WCA 使大多数的复制任务自动进行。WCA 为 WCA 复制的所有 WebSphere Commerce 表提供预定义的复制设置脚本文件。WCA 在运行脚本前将所有 WebSphere Commerce 表转换到您自己的环境中。

复制定制任务包含以下：

- 创建新的复制表
- 增加日志空间
- 计算表空间大小

- 修改复制控制表
- 启动“捕获”过程
- 注册复制过程
- 运行复制步骤
- 调度复制和抽取

创建新的复制表

要创建新的复制数据源表，请执行以下步骤：

1. 使用 DB2 控制中心（定义为复制源）或 DB2 DJRA 工具为 WCA 数据库中的表注册复制。
2. 通过执行以下步骤在 DB2 数据仓库中心创建复制步骤：
 - a. 在“复制属性”窗口中创建登台表。
 - b. 将具有整数类型的 **Load-Status** 列添加至登台表。
3. 将新的复制步骤链接至 DB2 仓库中心中的其它复制步骤。
4. 将新步骤提升为生产方式。

增加日志空间

复制源表所需的日志空间的增加取决于定义的数据源数目、复制源的行的长度、对那些表的更改次数以及要更新的列数。通常，日志空间应至少比这些表所需的原日志空间大三倍。

计算表空间大小

rep.sql 文件为更改数据（CD）表分配表空间，这些表用于在复制期间临时存储数据。要确定更改数据（CD）表的最小大小，请使用以下公式：

$$\begin{aligned} \text{minimum_CD_size} = & \\ & ((21 \text{ bytes}) + \text{sum}(\text{length of all registered columns})) * \\ & (\text{number of inserts, updates, and deletes to source table}) * \\ & (\text{exception factor}) \end{aligned}$$

要确定表的表空间大小，请使用以下公式：

$$\text{Table space size} = \text{minimum_CD_size} \div 4096$$

注：

1. 在计算复制数据的字节数时（公式的第一行），包含“捕获”程序添加至 CD 表的顶部数据的 21 字节。
2. 要确定对源表的插入、更新和删除次数（公式的第二行），请使用以下信息：

在估计 CD 表的空间要求时，无需为与 CD 表相关联的源表中存在的所有记录提供空间。第一次发生复制时，数据会绕过 CD 表导出至应用端的 **_R** 表。因此，CD 表仅需为估计将在复制运行之间的期间内更新、插入或删除的记录数提供足够空间。

要确定记录数，请使用以下公式：

$$\begin{aligned} \text{number of records updated} = & \\ & (\text{average number of rows}) * \end{aligned}$$

(factor for inserts/updates/deletes) *
(number of days capture data is held)

其中:

- (average number of rows) 可使用类似以下示例的 SQL 来计算:
(count(*)/count(distinct date(lastupdate))) from orders
- 对于 (factor for inserts/updates/deletes), 根据 WebSphere Commerce 发布更新、删除和插入的方式, 使用在 6 和 10 之间的因子。(该值是估计值。)
- (number of days capture data is held) 是 CD 表数据将保留的天数。

例如, 如果“捕获”程序每日一次修剪 CD 表中应用的行, 则间隔是 24 小时。如果 CD 表中的行有 100 字节(加上顶部的 21 字节)长, 则在 24 小时周期内可更新 100,000 条记录, 而 CD 表所需存储量约为 12 兆字节。(该示例使用异常因子 1。)

下表显示复制源表的表空间大小, 是使用时在时间间隔期间应用 1000 次更新这个值以及异常因子 1 来计算的:

表 6. 复制源表的表空间大小

WebSphere Commerce 源表	CD 表	WebSphere Commerce 表空间	CD 表大小 (包括 21 个字节)	最小 CD 大小	表空间 (CD 大小 ÷ 4096)
ADDRESS	CD_ADDRESS	TSADDRESS	3,221	3,221,000	787
ATTRIBUTE	CD_ATTRIBUTE	TSATTRIBUTE	1,145	1,145,000	280
CAMPAIGN	CD_CAMPAIGN	TSCAMPAIGN	1,285	1,285,000	314
CATALOG	CD_CATALOG	TSCATALOG	545	545,000	134
CATALOGDSC	CD_CATALOGDSC	TSCATALOGDSC	33,749	33,749,000	8,240
CATENTDESC	CD_CATENTDESC	TSCATENTDESC	131,677	131,677,000	32,148
CATENTREL	CD_CATENTREL	TSCATENTREL	617	617,000	151
CATENTRY	CD_CATENTRY	TSCATENTRY	1125	1,125,000	275
CATENTSHIP	CD_CATENTSHIP	TSCATENTSHIP	125	125,000	31
CATGPENREL	CD_CATGPENREL	TSCATGPENREL	307	307,000	75
CATGRPDESC	CD_CATGRPDESC	TSCATGRPDESC	34,007	34,007,000	8,303
CATGRPREL	CD_CATGRPREL	TSCATGRPREL	307	307,000	75
CATTOGRP	CD_CATTOGRP	TSCATTOGRP	53	53,000	13
CMPGNINTV	CD_CMPGNINTV	TSCMPGNINTV	29	29,000	8
CPGNLOG	CD_CPGNLOG	TSCPGNLOG	813	813,000	199
CPGNSTATS	CD_CPGNSTATS	TSCPGNSTATS	237	237,000	58
CURCONVERT	CD_CURCONVERT	TSCURCONVERT	69	69,000	17
FFMCENTDS	CD_FFMCENTDS	TSFFMCENTDS	32,813	32,813,000	8,011
FFMCENTER	CD_FFMCENTER	TSFFMCENTER	287	287,000	71
ICCNCPNLNK	CD_ICCNCPNLNK	TSICCNCPNLNK	293	293,000	72
ICCONSTRNT	CD_ICCONSTRNT	TSICCONSTRNT	56	56,000	14
ICEXPLFEAT	CD_ICEXPLFEAT	TSICEXPLFEAT	175	175,000	43
ICFEATPROP	CD_ICFEATPROP	TSICFEATROP	351	351,000	86

表 6. 复制源表的表空间大小 (续)

WebSphere Commerce 源 表	CD 表	WebSphere Commerce 表空 间	CD 表大 小 (包括 21 个字 节)	最小 CD 大小	表空间 (CD 大 小 ÷ 4096)
ICKNOWDESC	CD_ICKNOWDESC	TSICKNOWDESC	1053	1,053,000	258
ICKNOWLEDG	CD_ICKNOWLEDG	TSICKNOWLEDG	153	153,000	38
ICMETAPHOR	CD_ICMETAPHOR	TSICMETAPHOR	435	435,000	107
ICMETAREG	CD_ICMETAREG	TSICMETAREG	229	229,000	56
ICMREGDESC	CD_ICMREGDESC	TSICMREGDESC	257	257,000	63
IITEM	CD_IITEM	TSIITEM	353	353,000	87
INITIATIVE	CD_INITIATIVE	TSINITIATIVE	33,675	33,675,000	8,222
INTVMPE	CD_INTVMPE	TSINTVMPE	33	33,000	9
LISTPRICE	CD_LISTPRICE	TSLISTPRICE	116	116,000	29
MATYPE	CD_MATYPE	TSMATYPE	605	605,000	148
MPE	CD_MPE	TSMPE	351	351,000	86
MPETYPE	CD_MPETYPE	TSMPEETYPE	343	343,000	84
ORDERITEMS	CD_ORDERITEMS	TSORDERITEMS	1,179	1,179,000	288
ORDERS	CD_ORDERS	TSORDERS	751	751,000	184
ORDSTAT	CD_ORDSTAT	TSORDSTAT	872	872,000	213
PASTATS	CD_PASTATS	TSPASTATS	41	41,000	11
PCSTATS	CD_PCSTATS	TSPCSTATS	41	41,000	11
PESTATS	CD_PESTATS	TSPESTATS	105	105,000	26
SASTATS	CD_SASTATS	TSSASTATS	45	45,000	11
SHIPMODE	CD_SHIPMODE	TSSHIPMODE	6,151	615,000	151
STADDRESS	CD_STADDRESS	TSSTADDRESS	2,759	2,759,000	674
STORE	CD_STORE	TSSTORE	853	853,000	209
STORENTDS	CD_STORENTDS	TSSTORENTDS	32,817	32,817,000	8,012
STORELANG	CD_STORELANG	TSSTORELANG	32	32,000	8
STORETRANS	CD_STORETRANS	TSSTORETRANS	33	33,000	9
USERDEMO	CD_USERDEMO	TSUSERDEMO	764	764,000	187
USERS	CD_USERS	TSUSERS	1,864	1,864,000	456
USRTRAFFIC	CD_USRTRAFFIC	TSUSRTRAFFIC	99,343	99,343,000	24,254
WTAXINFO	CD_WTAXINFO	TSWTAXINFO	137	137,000	34
总计					103,330

修改复制控制表

在“源数据库的复制设置”配置步骤期间，在主管数据源的服务器上需要复制控制表和表空间。有关更多信息，请参阅《IBM WebSphere Commerce Analyzer V5.5 安装和配置指南》中的『配置复制』。

文件 %IWDA_DIR%\bin\db2\replication.bat 控制 WebSphere Commerce 服务器的复制设置。

在复制设置期间使用以下文件:

rep.sql

在 WebSphere Commerce 中创建复制控制表

rep_as400.sql

在 WebSphere Commerce (对于 **OS/400®**) 中创建复制控制表。

repContrlTables.sql

在 WCA 中创建复制控制表。

使用 **DJRA** 的“创建复制控制表”功能来创建复制控制表的脚本。

要更改 rep.sql 文件, 请按以下所示继续:

1. 在 bin\db2 目录中创建 replication.bat 文件的副本。
2. 编辑复制批处理文件:

- a. 找到以下行并将其注释掉。

```
@if exist %IWDA_DATA%\tmp\rep.sql del %IWDA_DATA%\tmp\rep.sql
```

- b. 找到以下行并将其注释掉。

```
@ %cmd% >%IWDA_DATA%\tmp\rep.sql 2>&1
```

重要事项: 输出文件必须命名为 rep.sql 且必须放在 %IWDA_DATA%\tmp。

3. 打开命令行窗口。
 - a. 将目录切换至您正在使用的数据源的子目录 (例如, 55be_ext)。
 - b. 输入以下命令:

```
rep TSPATH WCSSHEMA TS >%IWDA_DATA%\tmp\rep.sql
```

其中:

- *TSPATH* 是指向表空间位置的路径。

注: TSPATH 值必须以文件分隔符结束。

- *WCSSHEMA* 是 WebSphere Commerce 模式名称。
- *TS* 是缩放值。

4. 把目录更改为 %IWDA_DATA%\tmp。
5. 在开始“源数据库的复制设置”配置步骤之前按要求修改 rep.sql 脚本。

启动“捕获”过程

对于 WebSphere Commerce 服务器上的复制, 唯一必须作的任务是就在完成初始抽取之前启动“捕获”程序。请注意, 这是一次性的活动。

请参阅《*IBM WebSphere Commerce Analyzer V5.5 安装和配置指南*》中有关设置和启动“捕获”程序的过程。

注册复制过程

使用 IBM DB2 复制中心生成注册新表复制的脚本。WCA 配置管理器工具注册 WebSphere Commerce 数据库中的所有复制捕获表。

运行复制步骤

WCA 复制步骤是使用 IBM DB2 仓库中心复制工具设计的。复制步骤按链接顺序链接和处理。复制步骤将数据填充登台表，该表是 WebSphere Commerce 数据源的副本。然后，登台表使用 ETL（抽取、转换、移动和装入）步骤将数据填充 WCA 系统的数据集市。

开始复制步骤是 WCA ETL 处理的第一步。WCA 可以运行连续的复制处理也可以运行 ETL 处理。连续的复制处理避免了一次性移动大量数据，因为当数据开始可用时它会连续移动数据。

复制历史步骤提供信息以监视复制捕获服务器。当复制的最近同步时间早于先前的复制开始时间时，捕获服务器关闭，或在 WebSphere Commerce 系统中没有新数据。

WCA.PARAMETERS 表包含如下所示的复制信息：

replication_start

复制开始时间。此参数是控制步骤，帮助使复制处理保持连续。

replication_end

复制结束时间。此参数是控制步骤，帮助使复制处理保持连续。

replication_succeed

当复制开始时，设置为 1。当复制结束时，设置为 0。

request_etl_process

当复制步骤在处理时设置为 1，表明 ETL 处理应等待，直至所有复制步骤均得以处理。当 ETL 步骤开始时，设置为 0。如果 ETL 步骤在处理，复制处理将等待，直至所有 ETL 步骤均得以处理。如果没有 ETL 步骤在处理，则立即开始复制处理。此参数表明 ETL 处理将等待，直至完成所有复制步骤。

调度复制和抽取

如果没有正在处理的复制步骤，则对应于用户定义调度的 ETL 过程开始。可以将过程调度成按您希望的频率运行，但这些过程可能需要花费一些时间来运行。

第一次运行复制和抽取时，将复制并抽取所有现有数据。在后来的运行期间，仅复制和抽取新数据。

请参阅《*IBM WebSphere Commerce Analyzer V5.5 安装和配置指南*》中有关调度复制和抽取的指导。

第 13 章 定制数据挖掘

Professional Business WCA 数据挖掘组件处理数据挖掘模型的开发以及将其应用到 WCA 数据集市。要开发定制数据挖掘库以补充 WCA 提供的数据挖掘库，可以使用

Professional Business IBM DB2 Intelligent Miner for Data 工具。

数据挖掘概述

WCA 通过 **Professional Business** DB2 Intelligent Miner for Data 工具提供数据挖掘功能。DB2 Intelligent Miner for Data 工具增强了 WCA 服务器组织、聚集、总结和挖掘数据的能力。

下图显示 WCA 的数据挖掘功能：

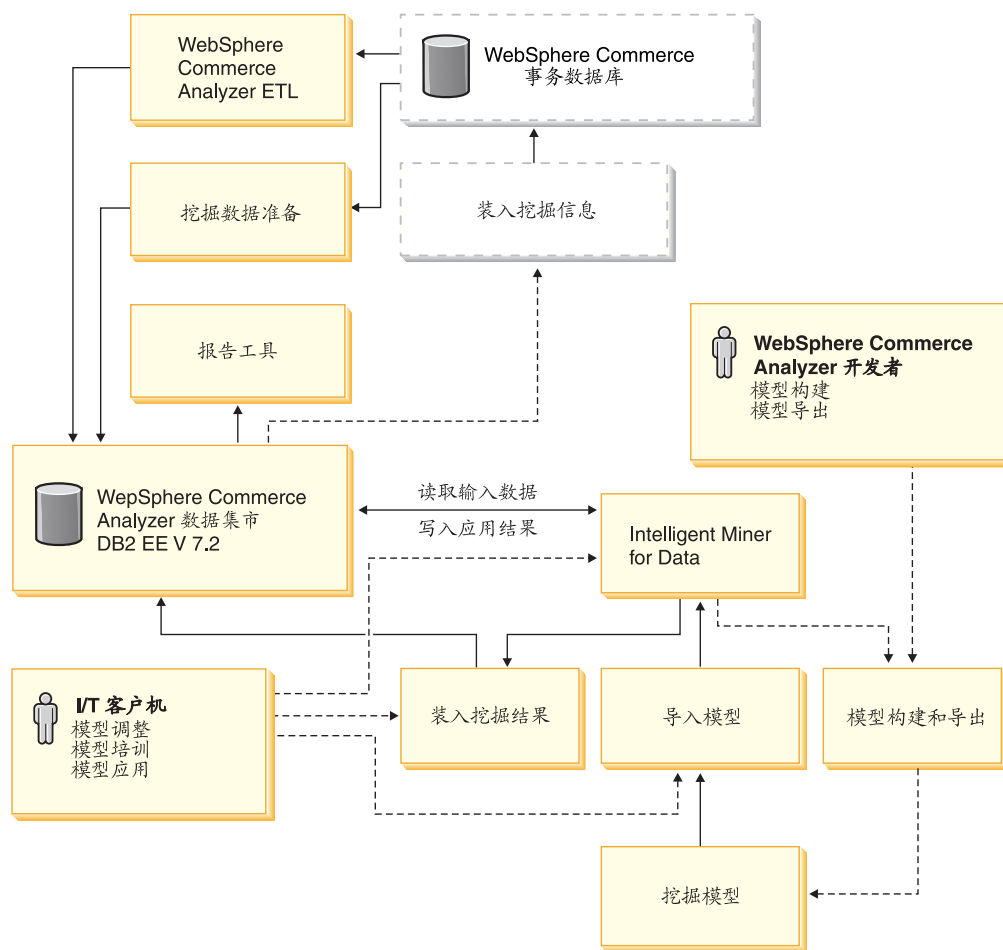


图 4. WCA 数据挖掘功能

WCA 提供以下数据挖掘功能：

- 支持应答一系列与数据挖掘相关的业务问题。WCA 根据这些问题提供挖掘模型。

- 在 WCA 的配置期间设置数据挖掘环境。该设置包括数据准备、表设置和挖掘模型设置。
- **Professional** **Business** 对用户通过 DB2 Intelligent Miner for Data GUI 执行数据挖掘操作的支持。这些数据挖掘操作从这些数据生成应答与数据挖掘相关的业务问题所必需的信息。
- 用于将 **Professional** **Business** DB2 Intelligent Miner for Data 挖掘结果装入到 WCA 数据集市中的实用程序。
- 合并数据挖掘结果的报表。

WCA 安装将在安装有 WCA 服务器的同一计算机上安装 **Professional** **Business** DB2 Intelligent Miner for Data 服务器。使用 WCA 模式在 WCA 数据集中创建与数据挖掘相关的数据库表。**Professional** **Business** DB2 Intelligent Miner for Data 从 WCA 数据集中读取输入数据。在运行数据挖掘序列后，使用与 WCA 一起提供的实用程序将数据挖掘结果装回到 WCA 数据集中。可以将这些数据挖掘结果合并到业务报表中。

用于群集的数据库模型示例

以下信息回答“购物者与订购总额之间有何关系”的问题：

输入数据表:	MEMBSUMS
样本数据集:	MEMBSUBSAMP
模型名称:	wcamemtotord61
结果文件:	wcamemtotord61.xml
应用输出表:	MEMAPPLY61
为模型指定的模型标识	90000-99999

在 DATATABS 表中注册数据表

data_id	data_name	data_type	usage_type	data_desc	data_location	sample_pct
1001	MEMBSUMS	0	0	成员使用活动的摘要表	wca	100
1002	MEMBAPPLY61	0	1	模型应用的输出表	wca	100
1101	wcamemtotord61.xml	2	1	来自 model_id 1051 的结果		
1011	MEMBSUMSAMP	0	1	MEMBSUMS 的抽样数据集	wca	50

在 **data_name** 和 **data location** 列中，必须在同一行上输入数据。如果此处的行中断，这是因为页面布局的局限。

在 MODELS 表中注册模型

model_id	model_name	model_type	model_alg	model_desc	data_id	file_id	res_file_name	file_location
1051	wcamem totord61	1	0	购物者与 订购总额 之间的关 系	1011	1101	memord val61.xml	

- 在 **model_name**、**res_file_name** 和 **file_location** 列中，必须在同一行上写数据。如果此处的行中断，这是因为页面布局的局限。
- **model_name** 应使用  IBM DB2 Intelligent Miner for Data 中定义的设置名称以用于模型培训。
- MODELS 表中的 **data_id** 和 **file_id** 指的是相应数据集在 DATATABS 表中的 **data_id**。

在 APPTABS 表中注册应用结果。

table_id	model_id	data_id	table_name	table_type	状态
1002	1051	1001	MEMAPPLY61	2	0


- APPTABS 模型中的 **model_id** 指的是所应用的模型在 MODELS 表中的 **model_id**。
- APPTABS 模型中的 **table_id** 和 **data_id** 指的是使用该模型的相应数据集在 DATATABS 表中的 **table_id** 和 **data_id**。

数据挖掘定制任务

数据挖掘定制任务包含以下：

- 创建数据挖掘模型
- 注册新数据挖掘模型
- 确定最近购买时间、购买频率和购买金额的值

创建数据挖掘模型

 可以使用 DB2 Intelligent Miner for Data 向导创建数据挖掘模型（又称为设置对象），并且可以使用设置笔记本对它们进行修改。

数据挖掘模型与相同挖掘库中其它对象具有内置关系。例如，挖掘设置对象需要您通过创建新数据对象或引用现有对象来指定输入数据。挖掘设置对象代表为挖掘功能指定的参数。这些参数之一是数据对象的名称，它用作输入数据的逻辑描述。

任何用户定义的输入或输出表均使用缺省数据挖掘表模式 **wcamng**。任何用户定义的用于数据挖掘的输入或输出表都应当使用前缀 **UX_**（对于用户扩展）。如果将模型培训的设置名称定义为 **ux_xxxx**，则应用此模型的相应设置将命名为 **ux_xxxxa**。

要创建新的数据挖掘模型，请参阅 *Using the Intelligent Miner for Data DB2* 文档。

注册新数据挖掘模型

标识在挖掘中使用的数据，并将其分组为一组表或视图。使用这些表和视图作为挖掘的输入数据。

在添加新的数据模型时，请在相应的控制表中注册它。必须使用下表：

- DATATABS
- MODELS
- APPTABS

DATATABS

输入和输出表的模型都在数据控制表中构建。在此表中将每个数据集注册为一个条目。

列：

列名	数据类型	注解
data_id	integer NOT NULL	已注册数据的唯一标识。
data_name	char(32) NOT NULL	数据集的名称。 <ul style="list-style-type: none">• 表或视图的名称（对于表或视图）。• 文件名称（对于文件）。
data_desc	varchar(254)	数据集的描述。
data_type	integer	数据集的类型。有效值是： 0 表 1 视图 2 文件
usage_type	integer	数据集的使用。有效值是： 0 输入 1 输出 2 输入 / 输出
app_type	integer	与数据集关联的应用程序类型。此列为可选列。
data_location	varchar(254) NOT NULL	以下值之一： <ul style="list-style-type: none">• 数据库名（对于表和视图）。• 全文件路径（对于文件）。
sample_pct	integer	数据集的样本大小（如果可用）。

索引：

名称	列名	类型
p_datatabs	data_id	主键

MODELS

MODELS 表管理关于模型、文件名和相关文件位置的信息。

列:

列名	数据类型	注解
model_id	integer NOT NULL	模型的唯一标识。
model_name	char(32)	模型的名称名称。
model_desc	varchar(254)	模型描述。
model_type	integer	模型类型。
model_alg	integer	指示用于构建模型的算法的标识。
data_id	integer	用于构建模型的数据的标识。
per_id	integer	最近培训模型的日期。
res_file_name	varchar(32)	包含预测模型标记语言 (PMML) 的文件名。
file_location	varchar(254)	保存文件的位置。

索引:

名称	列名	类型
p_resfiles	model_id	主键

APPTABS

APPTABS 表保存关于模型应用表的信息。每个模型应用都必须在此表中注册为一个条目。

列:

列名	数据类型	注解
table_id	integer NOT NULL	已注册表的唯一标识。
model_id	integer NOT NULL	模型应用的标识在模型表中注册。
data_id	integer	用于生成应用结果表的数据标识。
table_name	char(32) NOT NULL	保存应用结果的表的标识。
table_type	integer	输出表的主键。有效值是: 1 initiative_id 2 member_id 3 member_id 加 order_id 4 member_id 加 product_id 5 member_id 加 category_id 9 任何用户指定的表
状态	char(1)	指示表是否在动态使用的状态。

确定最近购买时间、购买频率和购买金额的值

最近购买时间、购买频率、购买金额（RFM）是根据变量来排序数据的方法，这些变量代表对商店的最近访问、所下的最大订单数目以及在订单上消费的最多金额。

在 WCA 中，最近购买时间、购买频率和购买金额（RFM）应用于与合同和帐户相关联的订单。以下条件确定 RFM:

- 购物活动的最近日期。
- 所下的最大订单数目。
- 订单消费的最多金额。

数据根据每个变量排序，然后划分为五个相等的容器。为每个合同或帐户指定 1 到 5 之间的数字，与其容器指定数字相对应。具备最近购物活动、最大订单数目和最多消费金额的合同或帐户的 RFM 值将为 555。具备最少活动、最小订单数目和最少消费金额的合同或帐户的 RFM 值将为 111。RFM 要求最少 5 条记录（容器的数量）以供成功运行。

附录 A. WCA 参数

此部分列出预定义的 WCA 参数。还讨论如何创建或更改 WCA 参数以及参数表在配置后可能看似如何。

预定义 WCA 参数的列表

参数影响数据集市在执行抽取后会看似如何。这些参数控制系统执行，也控制分析总结。使用 WCA 参数管理器修改参数。以下是 WCA.PARAMETERS 表的列名：

注：此参数列表在出版时为最新列表。

ROW_NUM

此参数提供系统生成的标识，与 PARAM_TYPE 一起用作 WCA.PARAMETERS 表的主键。

PARAM_TYPE

此参数是 WCA 定义的由抽取步骤、复制步骤和第三方报表工具使用的参数。一个或多个参数使您可定制缺省抽取的行为。如果 PARAM_TYPE 具有一个以上的 PARAM_VALUE，则在独立的行中表示每个 PARAM_VALUE。ROW_NUM 列区分每个 PARAM_TYPE 的多个值。例如：

表 7. 参数表示例

ROW_NUM	PARAM_TYPE	PARAM_VALUE
1	COMPLETED_STATUS	已完成
2	COMPLETED_STATUS	已装运

此表显示 PARAM_TYPE COMPLETED_STATUS 有两个值：“已完成”和“已装运”。

PARAM_VALUE

此参数是参数类型的字符串表示。对于整数，使用强制转型语句。

以下参数是当前为 WCA 数据集市定义的参数：

ABANDONED_MINUTES

此参数定义在成员更新未决订单后为放弃该订单所必须经历的时间段。参数 ABANDONED_ORD_STATUS 定义视为未决的订单状态值。此参数影响 FACT_ORDERS.ABANDONED 列以及 FACT_ORDERITEMS.ABANDONED 列。

将订单视为已放弃的缺省值是 **60**。

ABANDONED_ORD_STATUS

此参数定义 WebSphere Commerce 数据库的一系列视为未决订单的 ORDER.STATUS 和 ORDERITEMS.STATUS 值。使用此参数确定放弃订单还是订购商品。此参数影响 FACT_ORDERS.ABANDONED 列和 FACT_ORDERITEMS.ABANDONED 列。

以下是缺省值：

- A = 需要复查
- E = CSR 编辑
- I = 已提交订单

- L = 低库存
- M = 支付已开始
- P = 未决订单
- W = 等待核准
- X = 已取消

ANALYSIS_CATALOG_ID

对于每个产品目录标识，此参数占用一行。它为报表工具提供 WebSphere Commerce 中创建的产品目录的产品目录标识，以便跟踪类别中的度量。此产品目录必须为具有从产品目录到类别、类别到子类别以及子类别到产品之间的一对多关系的树。可以将此类型的分析称为产品目录的分类分析。此参数可有多个值。

BROKER_STORE_ID

此参数指示分类为“代理商店”的商店的 STORE_ID。在“选择网上商店和报表语言与货币”配置步骤期间进行该选择。每个代理商店对应一行。请参阅《IBM WebSphere Commerce Analyzer V5.5 安装和配置指南》中的『选择商店和语言』。

CONTRACT_ACTIVE

此参数确定 WCA.CONTRACT 表中指示合同仍然活动的 CON_STATUS_ID。状态代码的描述在 WCA.CON_STATUS_REF 表中。缺省值是 **3**。此参数可以有多个值。

CONTRACT_CANCELLED

此参数确定 WCA.CONTRACT 表中代表已取消合同的 CON_STATUS_ID。可在 WCA.CON_STATUS_REF 表中找到状态代码的描述。缺省值是 **5**。此参数可以有多个值。

CONTRACT_IN_PREPARATION

此参数确定 WCA.CONTRACT 表中指示合同仍在准备中的 CON_STATUS_ID。状态代码的描述在 WCA.CON_STATUS_REF 表中。缺省值是 **0**。此参数可以有多个值。

CPN_REDEEMED_ORDER_STATUS

缺省值为:

- 4 = 支付已开始
- 6 = 支付已授权
- 7 = 订单已发货

DMT_PROSPECT_ORD_STATUS

此参数为 WebSphere Commerce 数据库中的 ORDERS.STATUS 和 ORDERITEMS.STATUS 定义一系列的值。当成员拥有处于这些状态中任一状态的订单，而您尚未将该成员视为购买者，则可将该成员视为潜在购买者。此参数影响 WCA.MEMBER.DR_MEMBER_TYPE_ID 的值。

以下是缺省值:

- A = 需要复查
- E = CSR 编辑
- I = 已提交订单
- M = 支付已开始
- P = 未决订单
- Q = 订单模板
- W = 等待核准
- X = 已取消
- Y = 专用需求列表

Z = 可共享需求列表

DMT_PURCHASER_ORD_STATUS

此参数为 WebSphere Commerce 数据库中的 `ORDERS.STATUS` 和 `ORDERITEMS.STATUS` 定义一系列的值。当成员拥有处于这些状态中任一状态的订单时，该成员将被视为购买者。每个 `ORDER.STATUS` 值对应一行。在缺省情况下，WCA 将购买者定义为至少完成一份订单的用户。此参数影响 `WCA.MEMBER.DR_MEMBER_TYPE_ID` 的值。

以下是缺省值:

- B = 已延迟交货
- C = 支付已授权
- D = 已付押金
- F = 可用于远程供货
- G = 等待远程供货
- L = 低库存
- R = 可用于远程供货
- S = 订单已发货

EXTRACTION_COUNT

此参数定义已运行的抽取数。此参数为 IBM 保留参数。不要更改它。

EXTRACTION_SUCCEED

此参数指示抽取过程是否已成功完成。如果抽取过程成功完成，则此参数具有值 **0**。如果该过程不成功，则该参数具有值 **1**。此参数为 IBM 保留参数。不要更改它。

FE_EFFECTIVE_MINUTES

此参数定义“推荐销售”活动可能在其中影响买方对购买某一产品的决定的时间段。如果成员没有在定义的时间段内将广告产品放入购物车，则即使他以后购买了该产品，此活动也没有资格要求得到该产品的收入。

此参数影响下面 `WCA.FACT_EVENT` 列的填充:

- `SLS_VAL_CLKS`
- `STR_SLS_VAL_CLKS`
- `RPT_SLS_VAL_CLKS`
- `NUM_CKS_TO_ORD`

时间段的缺省值为 **60** 分钟。

FE_EFFECTIVE_ORD_STATUS

此参数为 WebSphere Commerce 数据库中的 `ORDERS.STATUS` 和 `ORDERITEMS.STATUS` 定义一系列的值。当“推荐销售”活动在 `FE_EFFECTIVE_MINUTES` 参数定义的时间段内向特定成员为特定产品作广告时，该成员将该产品放置到购物车中（未决订单），并随后购买了该产品，则该活动就可以要求负责此产品产生的收入。对于同一订购商品，可能会有许多适合的活动。该商品只能参加其中一个活动。

此参数影响下面 `WCA.FACT_EVENT` 列的填充:

- `SLS_VAL_CLKS`
- `STR_SLS_VAL_CLKS`
- `RPT_SLS_VAL_CLKS`
- `NUM_CKS_TO_ORD`

以下是缺省值:

- B = 已延迟交货
- C = 支付已授权
- D = 已付押金
- F = 可用于远程供货
- G = 等待远程供货
- L = 低库存
- R = 可用于远程供货
- S = 订单已发货

FM_EFFECTIVE_MINUTES

此参数定义在用户使用隐喻之时与用户购买该隐喻推动的产品且该隐喻仍可因该购买而得分之时的期间所允许经历的分钟数。使用此参数在 WCA.FACT_METAPHOR 表中填充销售额列。您可以修改此参数。缺省值为 **60** 分钟。

FM_EFFECTIVE_ORD_STATUS

此参数为 WebSphere Commerce 数据库中的 ORDERS.STATUS 和 ORDERITEMS.STATUS 定义一系列的值。当用户使用某产品隐喻，且用户在 FM_EFFECTIVE_MINUTES 指定的时间段内将产品放在其购物车中，则承认该隐喻促成了该产品的销售。

WCA.FACT_METAPHOR.STR_SLS_VAL_META 和

WCA.FACT_METAPHOR.RPT_SLS_VAL_META 列跟踪隐喻生成的收入。可以使用许多隐喻，但仅有将产品显示给用户的最近一个隐喻得分。所使用的值必须表明收入已实现。

以下是缺省值：

- B = 已延迟交货
- C = 支付已授权
- D = 已付押金
- F = 可用于远程供货
- G = 等待远程供货
- L = 低库存
- R = 可用于远程供货
- S = 订单已发货

MEMBER_ADDRESS_TYPE_ID

此参数帮助 WCA 决定用于在 MEMBER 列中填充基于地址的列的地址记录。可以将此参数与 ADDRESS 表中的 ADDRESS_TYPE_ID 字段比较。

MINING_APPLY_INTERVAL

此参数定义执行挖掘应用过程的频率。时间间隔表示为天数。

MINING_APPLY_TIME

此参数定义挖掘应用成功运行的时间戳记。如果挖掘应用操作成功，此时间戳记将具有此抽取运行的 TIME_CUT_OFF_LOCAL 值。此参数为 IBM 保留参数。不要更改它。

MINING_BASE_NAME

此参数定义具有预定义挖掘模型的挖掘库的名称。可以使用配置管理器更改此参数。

MINING_PASSWORD

此参数定义运行挖掘操作的用户的密码。可以使用配置管理器更改此参数。

MINING_TRACE

此参数有以下值:

- 0 = 最小跟踪级别
- 1 = 完全跟踪

0 是缺省值。

MINING_TRAINING_INTERVAL

此参数定义执行挖掘培训过程的频率。时间间隔表示为天数。

MINING_TRAINING_TIME

此参数定义挖掘培训成功运行的时间戳记。如果挖掘培训操作成功，此时间戳记将具有此抽取运行的 TIME_CUT_OFF_LOCAL 值。此参数为 IBM 保留参数。

MINING_USER_NAME

此参数定义运行挖掘操作的用户的名称。

NF_EFFECTIVE_ORDER_STATUS_ID

此参数创建与 FACT_ORDERITEMS 表中 ORDER_STATUS_ID 相对应的一系列 ORDER_STATUS_ID，代表作为已实现收入的订单一部分的订购商品条目。例如已装运、已完成或已延迟交货的订单。

以下是缺省值:

- 5 = 低库存
- 6 = 支付授权
- 7 = 订单已发货
- 13 = 已延迟交货
- 14 = 已针对供货发货
- 15 = 已付押金
- 17 = 供货已就绪
- 18 = 等待远程供货

NF_EFFECTIVE_MINUTES

此参数决定从被跟踪事件发生之时直至创建已完成订单之时所允许经历的分钟数。如果在此有效时间后创建订单，则该订单不与该特殊事件相关联。缺省值为 **60**

NON_PURGE_ORD_STATUS

此参数确定在从 ORDER 和 ORDERITEMS 表中删除记录后如何处理它们。如果删除的记录具有带 NON_PURGE_ORD_STATUS 标志的状态值，则 WCA 在相应的 WCA (FACT_ORDERS 或 FACT_ORDERITEMS 表) 中将 DELETED_STATUS 设置为 1。

此参数告知 WCA 哪些类型的记录可以作为用户操作的结果而非清除数据集市的结果而删除。作为用户操作结果而删除的记录具有设置为 1 的 DELETED_STATUS。作为数据集市清除结果而删除的记录继续存在于 WCA 数据集市中，同时其 DELETED_STATUS 设置为 0。

以下是缺省值:

- A = 需要复查
- E = CSR 编辑
- I = 已提交订单
- M = 支付已开始

- P = 未决订单
- Q = 订单模板
- X = 已取消
- Y = 专用需求列表
- Z = 可共享需求列表

ORDER_STATUS_BILLED

此参数确定 FACT_ORDERS 和 FACT_ORDERITEMS 表中代表已开票订单的 ORDER_STATUS_ID。可在 WCA.ORDER_STATUS_REF 表中找到状态代码的描述。

以下是缺省值:

- 7 = 订单已发货
- 14 = 已针对供货发货
- 15 = 已付押金
- 18 = 等待远程供货

ORDER_STATUS_CANCELLED

此参数确定 FACT_ORDERS 和 FACT_ORDERITEMS 表中代表已取消订单的 ORDER_STATUS_ID。状态代码的描述在 WCA.ORDER_STATUS_REF 表中。缺省值是 **5**。此参数可以有多个值。

ORDER_STATUS_COLLECTED

此参数确定 FACT_ORDERS 和 FACT_ORDERITEMS 表中代表已收帐订单的 ORDER_STATUS_ID。可在 WCA.ORDER_STATUS_REF 表中找到状态代码的描述。缺省值为 **15**。此参数可有多个值。

ORDER_STATUS_ID_NOEV

此参数帮助 WCA 确定 FACT_ORDERS 和 FACT_ORDERITEMS 表中代表可能未填充其货币金额列的订单。在这种情况下, WCA 将货币金额设置为 0, 将货币描述列设置为空。该值必须为整数。可以将该值与 ORDER_STATUS_ID 字段比较。缺省值为 **19**。此参数可有多个值。

ORDER_STATUS_ID_SUM_MEMBER

此参数确定 FACT_ORDERITEMS 表中的行必须与之相匹配的一系列 ORDER_STATUS_ID 值, 从而填充 SUM_MEMBER 表中的行。

以下是缺省值:

- 5 = 低库存
- 6 = 支付已授权
- 7 = 订单已发货
- 13 = 已延迟交货
- 14 = 已针对供货发货
- 15 = 已付押金
- 17 = 供货已就绪
- 18 = 等待远程供货

ORDER_STATUS_ID_SUM_TRADING

此参数确定 FACT_ORDERITEMS 表中的行必须与之相匹配的一系列 ORDER_STATUS_ID 值, 从而填充 SUM_TRADING 表中的行。

以下是缺省值:

- 5 = 低库存
- 6 = 支付已授权

- 7 = 订单已发货
- 13 = 已延迟交货
- 14 = 已针对供货发货
- 15 = 已付押金
- 17 = 供货已就绪
- 18 = 等待远程供货

ORDER_STATUS_XFERRED

缺省值为:

- R = 可用于远程供货
- G = 等待远程供货
- F = 可用于远程供货

ORDERS_AWAITING_PAYMENT

此参数帮助 WCA 确定哪些订单和订购商品仍在等待支付。可以将此参数与 FACT_ORDERS 和 FACT_ORDERITEMS 表中的 ORDER_STATUS_ID 进行比较。报表可使用此参数显示仅适用于等待支付的订单的度量。缺省值是 4。此参数可以有多个值。

ORG_BUSINESS_TYPE

此参数仅适用于商家到商家的交易。缺省值为 **OrgEntityBusinessType**。此参数确定 WCA 如何填充 WCA.ORGANIZATION.BUSINESS_TYPE 列。WCA 对 WebSphere Commerce MBRATTR 表使用此字符串以确定对 WebSphere Commerce MBRATTRVAL 表使用的 MBRATTR_ID，从而找到要放入 WCA.ORGANIZATION.BUSINESS_TYPE 列的字符串。

ORG_INDUSTRY_TYPE

此参数仅适用于商家到商家的交易。缺省值为 **OrgEntityIndustryType**。此参数确定 WCA 如何填充 WCA.ORGANIZATION.INDUSTRY_TYPE 列。WCA 对 WebSphere Commerce MBRATTR 表使用此字符串以确定对 WebSphere Commerce MBRATTRVAL 表使用的 MBRATTR_ID，从而找到要放入 WCA.ORGANIZATION.INDUSTRY_TYPE 列的字符串。

PRODUCT_PRICE_AGGREGATE

此参数同时适用于商家到商家和商家到消费者的交易。缺省值为 **MIN**，但有三个可能的值: MIN、MAX 和 AVG。此参数控制可确定 WCA 数据库中以下列的方法。

表 8. 交易方法

交易类型:	描述:
商家到商家	因为对单个商品可以有以 WCA.OFFER_PRICE 表中描述的多种货币表示的多种价格，因此需要使用聚集函数对每个产品仅选择一个估计价。此参数允许客户根据报表货币在此列中选择使用最低、最高或平均价格。
列:	
WCA.PRODUCT.RPT_EST_LIST_PRICE	
WCA.OFFER.EST_MQ_VALUE	

表 8. 交易方法 (续)

交易类型:	描述:
商家到消费者	在 WebSphere Commerce V5.5 中, 源是 WebSphere Commerce LISTPRICE 表。
列:	因为对某产品可有多种报价, 所以此参数允许客户根据要用于此列的报表货币向此列中填充最低、最高或平均价格。在 WebSphere Commerce Professional Edition V5.5 中, 源是 WebSphere Commerce OFFERPRICE 表。
WCA.PRODUCT.RPT_EST_LIST_PRICE	因为对单个商品可以有以 WebSphere Commerce OFFERPRICE 表中描述的多种货币表示的多种价格, 因此需要使用聚集函数对每个产品仅选择一个估计价。此参数允许客户根据要在此列中使用的报表货币选择最低、最高或平均价格。

REPORT_CURRENCY

此参数定义用于转换所有 RPT_XXXX 列的货币。通过使用 WCA 配置工具来选择该货币。所有 WCA 为之生成报表的商店必须可将其所有支持货币转换为该报表货币。

REPORT_LANGUAGE

此参数定义所有报表使用的语言。报表工具使用此参数确定如何显示特定于语言的名称和描述。此参数主要定义用于 XXXX_REF 表的公共语言。WCA 配置程序确定该值, 并且该值取决于当前安装的操作系统的语言。

REPLICATION_METHOD

此参数控制复制过程。如果该参数设置为 Y, 则使用连续复制。如果该参数设置为 N, 则使用调度复制。有关复制方法的讨论, 请参阅《IBM WebSphere Commerce Analyzer V5.5 安装和配置指南》中的『复制选项』。

REPLICATION_START

此参数为 IBM 保留参数。不要更改它。

REPLICATION_SUCCEED

此参数为 IBM 保留参数。不要更改它。

REQUEST_ETL_PROCESS

此参数为 IBM 保留参数。不要更改它。

RFM_BINS

此参数仅适用于商家到商家的交易。客户定义字段, 指定要用于划分数据的容器数目。缺省值为 5。在将记录排序且等量划分入 5 个容器后, 会根据与每个记录相关联的容器为其指定相应的 1、2、3、4 或 5 的 R、F 或 M 值。复合值是从 111 到 555 的所有变体。该字段的理想值是 3、4 或 5。

RFM_INTERVAL

此参数是可定义的字段, 确定运行最近购买时间、购买频率和购买金额 (RFM) 的频率 (以天计)。此参数为 IBM 保留参数。不要更改它。

RFM_LAST_RUN

此参数指示上次运行 RFM 的时间。此参数为 IBM 保留参数。不要更改它。

RFQRSP_OUTSTANDING_ORDERS

此参数确定 FACT_ORDERS 和 FACT_ORDERITEMS 表中代表未处理订单的 ORDER_STATUS_ID。可以将与带有这些标志的 RFQRSPY 相关的订单视为未处理，因为客户尚未启动支付。状态代码的描述在 WCA.ORDER_STATUS_REF 表中。

以下是缺省值:

- 1 = 未决订单
- 3 = 已提交订单
- 4 = 支付已开始
- 9 = 等待核准
- 11 = 需要复查
- 12 = CSR 编辑

RFQ_RESPONSE_IN_PREPARATION

此参数帮助 WCA 确定哪些 RFQ_RSP 记录仍处于准备状态。可以将此多值参数与 RFQ_RSP.RSP_STATUS_ID 字段比较。缺省值为 1。

RFQ_WINNING_RESPONSES

此参数确定哪些 RSP_STATUS_ID 代表作为中标响应的 RFQ 响应。状态代码的描述在 WCA.RSP_STATUS_REF 表中。

以下是缺省值:

- 8 = 中标
- 10 = 中标完成

SRF_EFFECTIVE_ORDER_STATUS_ID

此参数创建与 FACT_ORDERITEMS 表中 ORDER_STATUS_ID 相对应的一系列 ORDER_STATUS_ID，代表作为已实现收入的订单一部分的订购商品条目。例如已装运、已完成或已延迟交货的订单。

以下是缺省值:

- 5 = 低库存
- 6 = 支付已授权
- 7 = 订单已发货
- 13 = 已延迟交货
- 14 = 已针对供货发货
- 15 = 已付押金
- 17 = 供货已就绪
- 18 = 等待远程供货

SRF_EFFECTIVE_MINUTES

此参数决定从被跟踪事件发生之时直至创建已完成订单之时所允许经历的分钟数。如果在此有效时间后创建订单，则该订单不与该特殊事件相关联。缺省值为 60

STORE

此参数指示所有在此 WCA 实例期间生成报表的商店的 STORE_ID。配置工具设置此参数。每个商店占用一行。

TIME_CUT_OFF

此参数定义抽取窗口的结束，这是在抽取和复制过程开始时设置的。WebSphere Commerce 服务器提供此参数。此参数为 IBM 保留参数。不要更改它。

TIME_CUT_OFF_LOCAL

此参数定义抽取窗口的结束。这是在抽取和复制过程开始时在相对于本地 WCA 数据库的时间戳记中设置的。此参数为 IBM 保留参数。不要更改它。

TIME_CUT_OFF_PREV

此参数定义抽取窗口的开始。这是在抽取和复制过程开始时设置的。WebSphere Commerce 服务器提供此参数。此参数为 IBM 保留参数。不要更改它。

WCS_SOURCE

此参数定义要从中抽取 WCA 数据的 WebSphere Commerce 数据库的版本。当前，此参数具有值 55be_ext。此参数为 IBM 保留参数。不要更改它。

下表是 IBM Tivoli Web Site Analyzer 列的列表：

wsa_category

缺省值为 catGroupId。

wsa_coupon

缺省值为 couponIds。

创建或更改参数

要快速访问参数，请使用 WCA 参数管理器。

下表包含参数管理器中的一系列条目，使您可访问用于 WCA 参数维护的面板：

表 9. WCA 参数和用于参数维护的面板

条目	WCA 参数
商店	store、REPORT_CURRENCY、 REPORT_LANGUAGE、BROKER_STORE_ID
产品目录	ANALYSIS_CATALOG_ID
挖掘模型	MINING_USER_NAME、MINING_APPLY_TIME、 MINING_BASE_NAME、MINING_PASSWORD、 MINING_TRACE、MINING_TRAINING_INTERVAL、 MINING_TRAINING_TIME、MINING_USER_ID
RFM	RFM_INTERVAL、RFM_BINS、RFM_LAST_RUN、 ORDER_STATUS_ID_SUM_MEMBER、 ORDER_STATUS_ID_SUM_TRADING、 ORDER_STATUS_XFERRED、RFM_LAST_RUN
已放弃的订单	ABANDONED_MINUTES、ABANDONED_ORD_STATUS
订单属性	NON_PURGE_ORD_STATUS、 ORDERS_AWAITING_PAYMENT、 ORDER_STATUS_BILLED、 ORDER_STATUS_CANCELLED、 ORDER_STATUS_COLLECTED、 ORDER_STATUS_ID_NOREV
成员属性	ADDRESS_MEMBER_TYPE_ID、 DMT_PROSPECT_ORD_STATUS、 DMT_PURCHASER_ORD_STATUS
合同	CONTRACT_ACTIVE、CONTRACT_IN_PREPARATION、 CONTRACT_CANCELLED

表 9. WCA 参数和用于参数维护的面板 (续)

RFQ 属性	RFQ_RESPONSE_IN_PREPARATION, RFQ_WINNING_RESPONSE, RFQ_OUTSTANDING_ORDER
财政阶段	
订单关联	FE_EFFECTIVE_MINUTES, FE_EFFECTIVE_ORD_STATUS, FM_EFFECTIVE_MINUTES, FM_EFFECTIVE_ORD_STATUS, CPN_REDEEMED_ORDER_STATUS
订单状态	
附加属性	PRODUCT_PRICE_AGGREGATE、ORG_BUSINESS_TYPE、 ORG_INDUSTRY_TYPE、HOTSPOT_LIMIT、 WSA_PARM_TYPE

在“维护 WCA 参数”窗口中，可以通过在**参数类型**字段中输入新参数名称来创建参数。正如对其它客户定义的步骤、函数和外部调用的脚本文件名称一样，请为客户定义的参数名称加前缀 *UX*，以便很容易将其标识为客户定义参数而非随 WCA 一起提供的参数。该 *UX* 前缀命名约定仅是一个建议。wcaparm.properties 位于 %IWDA_DIR%\lib。

可以在**参数值**字段中输入参数值或多个值，从而不必通过在属性文件中定义它来添加值。或者，如果希望为您自己的参数在一系列值中允许一个值或指定缺省值，请创建客户定义的 wcaparm.properties 文件。WCA 使用该属性文件中的定义来填充“维护 WCA 参数”窗口。

以下是 wcaparm.properties 文件的一些条目的示例:

```

parm1.name=UX_MyOwnParm parameter identification (PARAM_TYPE)
parm1.type=s          how is the parameter handled on the panel:
                        1=single value,
                        n=multiple values,
                        s=predefined values
parm1.values=Y;N      values for the drop-down selection box
parm1.defaults=Y      default setting
parm1.prog=generic

parm2.name=UX_MySecondParm
parm2.type=1
parm2.defaults=15
parm2.prog=generic
    
```

WCA 提供的 wcaparm.properties 文件位于 lib 子目录中。如果您为自己的参数创建另一 wcaparm.properties 文件，请将其放在 tmp 目录下。

配置后的参数表

以下是 WCA.PARAMETERS 表在配置后可能显示的样子（不带 **last_update** 和 **last_update_ID** 字段）。

REPORT_LANGUAGE 和 REPORT_CURRENCY 参数类型值起源于配置工具。
TIME_CUT_OFF、TIME_CUT_OFF_PREV、TIME_CUT_OFF_LOCAL、
EXTRACTION_COUNT 和 EXTRACTION_SUCCEED 参数类型值起源于抽取过程。

在此展示的 WCA.PARAMETERS 表中所显示的值是缺省值。此列表中的所有参数可以具有 WCA PARAMETER 表中的多个值。而且仅在 .sql 子句中使用这些参数。

ROW_NUM	PARAM_TYPE	PARAM_VALUE
1	ABANDONED_MINUTES	60
1	ABANDONED_ORD_STATUS	A
2	ABANDONED_ORD_STATUS	E
3	ABANDONED_ORD_STATUS	I
4	ABANDONED_ORD_STATUS	P
5	ABANDONED_ORD_STATUS	L
6	ABANDONED_ORD_STATUS	M
7	ABANDONED_ORD_STATUS	W
8	ABANDONED_ORD_STATUS	X
1	ANALYSIS_CATALOG_ID	20101000000
2	ANALYSIS_CATALOG_ID	21101000000
1	BROKER_STORE_ID	211
1	CONTRACT_ACTIVE	3
1	CONTRACT_CANCELLED	5
1	CONTRACT_IN_PREPARATION	0
1	CPN_REDEEMED_ORDER_STATUS	4
2	CPN_REDEEMED_ORDER_STATUS	6
3	CPN_REDEEMED_ORDER_STATUS	7
1	DMT_PROSPECT_ORD_STATUS	A
2	DMT_PROSPECT_ORD_STATUS	E
3	DMT_PROSPECT_ORD_STATUS	I
4	DMT_PROSPECT_ORD_STATUS	W
5	DMT_PROSPECT_ORD_STATUS	M
6	DMT_PROSPECT_ORD_STATUS	P
7	DMT_PROSPECT_ORD_STATUS	Q
8	DMT_PROSPECT_ORD_STATUS	X
9	DMT_PROSPECT_ORD_STATUS	Y
10	DMT_PROSPECT_ORD_STATUS	Z
1	DMT_PURCHASER_ORD_STATUS	B
2	DMT_PURCHASER_ORD_STATUS	C
3	DMT_PURCHASER_ORD_STATUS	D
4	DMT_PURCHASER_ORD_STATUS	F
5	DMT_PURCHASER_ORD_STATUS	G
6	DMT_PURCHASER_ORD_STATUS	L
7	DMT_PURCHASER_ORD_STATUS	R
8	DMT_PURCHASER_ORD_STATUS	S
1	EXTRACTION_COUNT	2
1	EXTRACTION_SUCCEED	0
1	EXTRACTION_TRACE	0

ROW_NUM	PARAM_TYPE	PARAM_VALUE
1	FE_EFFECTIVE_MINUTES	60
1	FE_EFFECTIVE_ORD_STATUS	B
2	FE_EFFECTIVE_ORD_STATUS	C
3	FE_EFFECTIVE_ORD_STATUS	D
4	FE_EFFECTIVE_ORD_STATUS	F
5	FE_EFFECTIVE_ORD_STATUS	G
6	FE_EFFECTIVE_ORD_STATUS	L
7	FE_EFFECTIVE_ORD_STATUS	R
8	FE_EFFECTIVE_ORD_STATUS	S
1	FISCAL_PERIODS_UNTIL	2005
1	FISCAL_YEAR_START	C0101
1	FM_EFFECTIVE_MINUTES	60
1	FM_EFFECTIVE_ORD_STATUS	B
2	FM_EFFECTIVE_ORD_STATUS	C
3	FM_EFFECTIVE_ORD_STATUS	D
4	FM_EFFECTIVE_ORD_STATUS	F
5	FM_EFFECTIVE_ORD_STATUS	G
6	FM_EFFECTIVE_ORD_STATUS	L
7	FM_EFFECTIVE_ORD_STATUS	R
8	FM_EFFECTIVE_ORD_STATUS	S
1	MEMBER_ADDRESS_TYPE_ID	3
1	MINING_APPLY_INTERVAL	1
1	MINING_APPLY_TIME	1000-01-01-00.00.00
1	MINING_BASE_NAME	mine0215
1	MINING_CLOSED_LOOP	Y
1	MINING_PASSWORD	<i>variable</i>
1	MINING_TRACE	0
1	MINING_TRAINING_INTERVAL	1
1	MINING_TRAINING_TIME	1000-01-01-00.00.00
1	MINING_USER_NAME	martuser
1	NF_EFFECTIVE_MINUTES	60
1	NF_EFFECTIVE_ORD_STATUS_ID	5
2	NF_EFFECTIVE_ORD_STATUS_ID	6
3	NF_EFFECTIVE_ORD_STATUS_ID	7
4	NF_EFFECTIVE_ORD_STATUS_ID	13
5	NF_EFFECTIVE_ORD_STATUS_ID	14
6	NF_EFFECTIVE_ORD_STATUS_ID	15
7	NF_EFFECTIVE_ORD_STATUS_ID	17
8	NF_EFFECTIVE_ORD_STATUS_ID	18
1	NON_PURGE_ORD_STATUS	A
2	NON_PURGE_ORD_STATUS	E

ROW_NUM	PARAM_TYPE	PARAM_VALUE
3	NON_PURGE_ORD_STATUS	I
4	NON_PURGE_ORD_STATUS	P
5	NON_PURGE_ORD_STATUS	L
6	NON_PURGE_ORD_STATUS	M
7	NON_PURGE_ORD_STATUS	Q
8	NON_PURGE_ORD_STATUS	X
9	NON_PURGE_ORD_STATUS	Y
10	NON_PURGE_ORD_STATUS	Z
1	ORDERS_AWAITING_PAYMENT	4
1	ORDER_STATUS_BILLED	7
2	ORDER_STATUS_BILLED	14
3	ORDER_STATUS_BILLED	15
4	ORDER_STATUS_BILLED	18
1	ORDER_STATUS_CANCELLED	2
1	ORDER_STATUS_COLLECTED	15
1	ORDER_STATUS_ID_NOREV	19
1	ORDER_STATUS_ID_SUM_MEMBER	5
2	ORDER_STATUS_ID_SUM_MEMBER	6
3	ORDER_STATUS_ID_SUM_MEMBER	7
4	ORDER_STATUS_ID_SUM_MEMBER	13
5	ORDER_STATUS_ID_SUM_MEMBER	14
6	ORDER_STATUS_ID_SUM_MEMBER	15
7	ORDER_STATUS_ID_SUM_MEMBER	17
8	ORDER_STATUS_ID_SUM_MEMBER	18
1	ORDER_STATUS_ID_SUM_TRADING	5
2	ORDER_STATUS_ID_SUM_TRADING	6
3	ORDER_STATUS_ID_SUM_TRADING	7
4	ORDER_STATUS_ID_SUM_TRADING	13
5	ORDER_STATUS_ID_SUM_TRADING	14
6	ORDER_STATUS_ID_SUM_TRADING	15
7	ORDER_STATUS_ID_SUM_TRADING	17
8	ORDER_STATUS_ID_SUM_TRADING	18
1	ORDER_STATUS_XFERRED	F
2	ORDER_STATUS_XFERRED	G
3	ORDER_STATUS_XFERRED	R
1	ORG_BUSINESS_TYPE	OrgEntityBusinessType
1	ORG_INDUSTRY_TYPE	OrgEntityIndustryType
1	PRODUCT_PRICE_AGGREGATE	MIN
1	REPLICATION_METHOD	N
1	REPLICATION_START	2003-02-17 16:26:19.359002

ROW_NUM	PARAM_TYPE	PARAM_VALUE
1	RELOCATION_SUCCEED	0
1	REPORT_CURRENCY	CAD
1	REPORT_LANGUAGE	- 1
1	REQUEST_ETL_PROCESS	0
1	RFM_BINS	5
1	RFM_INTERVAL	1
1	RFM_LAST_RUN	1000-01-01-00.00.00
1	RFQRSP_OUTSTANDING_ORDERS	11
2	RFQRSP_OUTSTANDING_ORDERS	12
3	RFQRSP_OUTSTANDING_ORDERS	3
4	RFQRSP_OUTSTANDING_ORDERS	1
5	RFQRSP_OUTSTANDING_ORDERS	4
6	RFQRSP_OUTSTANDING_ORDERS	9
1	RFQ_RESPONSE_IN_PREPARATION	1
1	RFQ_WINNING_RESPONSES	8
2	RFQ_WINNING_RESPONSES	10
1	SRF_EFFECTIVE_MINUTES	60
1	SRF_EFFECTIVE_ORD_STATUS_ID	5
2	SRF_EFFECTIVE_ORD_STATUS_ID	6
3	SRF_EFFECTIVE_ORD_STATUS_ID	7
4	SRF_EFFECTIVE_ORD_STATUS_ID	13
5	SRF_EFFECTIVE_ORD_STATUS_ID	14
6	SRF_EFFECTIVE_ORD_STATUS_ID	15
7	SRF_EFFECTIVE_ORD_STATUS_ID	17
8	SRF_EFFECTIVE_ORD_STATUS_ID	18
1	TIME_CUT_OFF	2003-02-17 16:26:19.359002
1	TIME_CUT_OFF_LOCAL	2003-02-17 16.51.21.109002
1	TIME_CUT_OFF_PREV	1000-01-01-00.00.00
1	WCS_SOURCE	55be_ext

附录 B. 声明

本信息是为在美国提供的产品和服务编写的。IBM 可能在其它国家或地区不提供本文中讨论的产品、服务或特征。有关当前在您所在区域可获取的产品和服务的信息，请向您的本地 IBM 代理咨询。任何对 IBM 产品、程序或服务的引用并不明示或默示只能使用 IBM 产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务都可以用来代替。但是，任何非 IBM 产品、程序或服务功能的评估和验证均由用户自行负责。

IBM 可能已经申请或正在申请与本文档有关的各项专利权。提供本文档并不表示允许您使用这些专利。您可以用书面方式将许可查询寄往：

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

要获取关于双字节 (DBCS) 许可证查询的信息，请与您所在国家或地区的 IBM 知识产权部门联系，或用书面方式将查询寄往：

IBM World Trade Asia Corporation Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japan

以下段落对英国或其它类似国家或地区不适用，因为在那些国家或地区中当地法律不允许供应这些产品：国际商业机器公司以“按现状”的基础提供本出版物，不附有任何形式的（无论是明示的还是默示的）保证，包括（但不限于）合法性、适销性和适用于某特定用途的默示保证。一些地区在某些交易中不允许放弃明示或默示的保证，因此本条款也不一定适用于您。

本信息中可能包含有技术方面不够准确的地方或印刷错误。此处的信息将定期更改；这些更改将包含在信息的新版本中。IBM 可以随时对本信息中描述的产品和 / 或程序进行改进和 / 或更改，而不另行通知。

本信息中对非 IBM Web 站点的引用都只是为了方便起见才提供的且不会在任何方式下对那些 Web 站点作保证。该 Web 站点中的资料不是 IBM 产品资料的一部分，使用那些 Web 站点带来的风险将由您自行承担。

IBM 可能会以它认为适当的任何方式来使用和发布您提供的任何信息而无须对您负责。

为了以下目的：(i) 允许在独立创建的程序和其它的程序（包括本程序）之间进行信息交换 (ii) 允许对已经交换的信息进行相互使用，而希望获取本程序有关信息的合法用户请与下列地址联系：

IBM Corporation
J46A/G4
555 Bailey Avenue
San Jose, CA 95141-1003
U.S.A.

只要遵守适当的条件和条款，包括某些情形下的一定数量的付款，都可获取这方面的信息。

按照 IBM 客户协议、国际程序许可证协议或任何等同协议中的条款，本文档中描述的许可程序及其所有可用的许可资料均由 IBM 提供。

此处包含的任何性能数据都是在受控环境中测得的。因此，在其它操作环境中获取的数据可能会有明显的不同。有些评测可能是在开发级的系统上进行的，因此无法保证在一般可用系统上进行的评测是相同的。此外，有些评测是通过推算而估测的。实际结果可能会有所不同。本文档的用户应当验证其特定环境的适用数据。

涉及非 IBM 产品的信息可从这些产品的供应商、其发布的公告或其它公众可用源处获取。IBM 没有对这些产品进行测试，也无法确认其性能的精确性、兼容性和其它任何非 IBM 产品相关的特性。有关非 IBM 产品性能的问题应当由这些产品的供应商负责。

对于有关 IBM 未来的方向或意向的声明，可以对其更改或取消而不另行通知，它们仅仅是目标和意愿。

显示的所有 IBM 价格是 IBM 当前的建议零售价，可以随时更改，而不另行通知。经销商价格可能是不同的。

此产品包含由 Apache Software Foundation 开发的软件 (<http://www.apache.org/>)。

商标

以下术语是国际商业机器公司在美国和 / 或其它国家或地区的商标:

DB2
DB2 Universal Database
IBM
Intelligent Miner
WebSphere

Microsoft、Windows、Windows NT[®] 和 Windows 徽标是 Microsoft Corporation 在美国和 / 或其它国家或地区的商标。

Java 和所有基于 Java 的商标和徽标是 Sun Microsystems, Inc. 在美国和 / 或其它国家或地区的商标或注册商标。

其它公司、产品和服务的名称可能是其它公司的商标或服务标记。

词汇表

[B]

报表应用程序 (reporting application)：用来搜集客户和业务销售交易有关信息的程序。

表 (table)：所命名的数据对象，由特定数量的列和一些未排序的行构成。

[C]

仓库中心控制数据库 (Warehouse Center Control Database)：包含存储仓库中心元数据所需控制表的仓库中心数据库。

产品比较 (Product Comparison)：“产品比较”是一种隐喻，允许用户比较两个或两个以上的产品。

产品探测 (Product Explorer)：“产品探测”是一种隐喻，允许用户设置产品的一些特征要求/约束（成本、颜色、类型等等），并搜索匹配产品。

抽取时间窗口 (extraction time window)：上一次为该源运行 WCA 抽取的时间与当前时间之间的时间窗口。对于 WebSphere Commerce 5.5 源，这是由 WCA 参数 TIME_CUT_OFF 和 TIME_CUT_OFF_PREV 指示的。

抽取 (extraction)：将数据从数据库抽出。对于 WCA，将数据从 WCA 服务器上的临时表移动至 WCA 数据集市的过程。临时表中的数据是从 WebSphere Commerce 数据库中复制的。

[D]

定制 (customization)：用户对 WCA 的补充或更改，以便更加符合单个电子交易的业务模型。

[F]

复制 (replication)：将所定义的数据集维护在多个位置的处理。它涉及将指定更改从某个位置复制到另一个位置以及使这两个位置中的数据同步。对于 WCA，将数据从 WebSphere Commerce 数据库中移动至 WCA 服务器上的临时表中的处理。

[G]

供货 (fulfillment)：接收订单时所发生的过程。供货过程通常包含诸如订单管理、装运管理、退货和状态跟踪之类的任务。

广告 (advertisement, ad)：一种为提高产品或服务知名度而发布或广播的营销素材。在 Web 上，最常见的广告类型是标题广告。

[H]

活动 (initiative)：一种应用技术，用以鼓励特定的行为，例如购买一样产品。

[J]

竞销 (campaign)：一系列有计划的运作，用以达到一系列原先定义好的业务目标。在零售市场营销中，活动是用于达到竞销目标的常用方法。

[K]

客户会话数据 (customer session data)：在客户访问网上商店期间从客户处所搜集的信息。

扩展 (extension)：用户对 WCA 组件的补充（例如，新的报表或新的数据集市表）。

[L]

列 (column)：关系数据库管理系统中属性的名称。构成对特定实体的描述的列值的集合称为行。一列等同于非关系文件系统中一条记录中的一个字段。

[S]

视图 (view)：一个或多个表中数据的一种可选表示。一个视图可包含定义它的一个或多个表中的全部或部分列。

数据集市 (datamart)：数据仓库的子集，它包含为满足某个部门或小组的特定需要而定制的数据。数据集市可以是您整个组织所用仓库的子集，例如包含在 OLAP 工具中的数据。

数据库管理存储器 (Database Managed

Storage, DMS): 一种数据存储类型, 其中的表空间由数据库管理员管理。在创建表时指定表空间大小并分配空间。

数据挖掘 (data mining): 从数据仓库中收集关键业务信息、使信息相互关联然后揭示关联性、模式和趋势的过程。

[W]

挖掘库 (mining base): 存储了所有关于挖掘运行设置和相应结果的信息的库。

[X]

系统管理存储器 (SMS, System Managed Storage): 一种类型的数据存储器, 其中操作系统管理受硬盘大小限制的表空间。数据随机存储在硬盘上表空间的目录容器 (文件系统中的目录名) 中。

销售助手 (Sales Assistant): “销售助手”是一种隐喻, 面向不熟悉产品详细信息且不会设置特征要求的购物者。

形象 (impression): 代表 Web 页面上竞销、活动和电子广告位的组合, 用以向客户提供信息并允许客户通过单击与此组合关联的链接来利用此信息。

修订 (revision): 用户对现有 WCA 组件的更改 (例如, 报表或数据集市表)。

[Y]

业务问题 (business question): 业务报表中所回答的有关成功举行各种竞销、活动以及正在访问商店的客户的特定信息的问题。

隐喻 (metaphor): WebSphere Commerce 的一种功能, 作为“产品顾问”组件的一部分而提供。它为购物者浏览产品提供了三个使用范例 (或隐喻): “产品探测”、“销售助手”和“产品比较”。“产品探测”隐喻允许用户设置产品的一些特征要求/约束 (成本、颜色、类型等等), 并搜索匹配产品。“销售助手”隐喻面向不熟悉产品详细信息且不会设置特征要求的购物者。此隐喻提出一系列问题, 然后从中推断客户想要的产品。“产品比较”隐喻允许用户比较两个或两个以上的产品。有关隐喻的更多信息, 请参阅 WebSphere Commerce 文档。

用户注册属性文件 (User Registration properties file): WebSphere Commerce 服务器上包含支持商店使用正确语言和国家或地区所需信息的文件。

E

ETL: 抽取、转换和装入。将数据从某个数据库抽出并将其放入另一个不同类型的数据库时所执行的功能。

O

ODBC 名称 (ODBC name): 数据库的“开放式数据库连接”名称。

ODS: 运作数据存储。用于 ETL 处理的工作区域。WebSphere Commerce 中的数据将复制到 ODS (_r 表)。

P

PMML: Predictive Model Markup Language (预测模型标记语言) 的缩写。一种由“数据挖掘组”定义、基于 XML 的语言, 它为公司提供方法来定义在认可供应商的应用程序之间的预测模型和共享模型。

R

RFM: 最近购买时间、购买频率、购买金额。一种技术, 通过检查客户最近购买的时间 (最近购买时间)、客户购买的频率 (购买频率) 以及客户消费的金额 (购买金额), 从而确定哪些客户是最佳客户。

RFQ: 报价请求。对供应商的投标邀请, 邀请其对提供公司或中介所需的描述产品或服务进行投标。

W

WebSphere Application Server: 一种全面的基于 Java 2 Platform Enterprise Edition (J2EE) 1.3 和 Web 服务技术的应用程序服务器, 它将企业数据和事务与电子商务世界集成起来。通过丰富的应用程序部署环境, 您可以构建、管理并部署动态电子商务应用程序; 可以处理大量事务; 还可以将后端业务数据和应用程序扩展至 Web。

索引

[B]

- 报表框架 5
- 报表应用程序
 - 描述 5
- 报表, 业务 7
- 报价请求 7
- 笔记本, 设置 67
- 编码标准 51, 58
- 标志
 - 工作日 18
 - 假日 18
- 表
 - 重新组织 23
 - 删除记录 20
 - LAST_UPDATED_REF 50
 - WCA.PARAMETERS 71, 81
- 表空间, 计算 59
- 表空间, 为 DMS 更改 19
- 表, 为 DMS 更改 19
- 捕获 57

[C]

- 财政年度修改, 进行 19
- 参考表 13
- 参数管理员
 - 定义 21
 - 使用“TraceLog 查看器”工具 21
- 参数值 71
- 插入 sql 语句 38
- 产品顾问 8
- 抽取
 - 定义 4
- 抽取过程
 - 编码标准 51
 - 参数 71
 - 配置后的参数 81
 - 疑难解答 13
- 抽取、转换、移动和装入
 - 请看 ETML
- 抽取, 定制 51
- 除去
 - 商店 13
 - 商店的事实和尺寸表 13
 - 数据集市 11
- 除去数据库 11
- 除去/删除 11
- 创建
 - 数据挖掘模型 67
- 存储, 除去 13

- 错误处理 40
- 错误码 52

[D]

- 登台表 58
- 调度复制和抽取 63
- 订单 8
- 定制
 - 规则和准则 49
 - 扩展 43
 - 修订 44
- 定制抽取 51
- 定制方案 44
- 定制规则 49
- 定制准则 49

[F]

- 复制
 - 定义 4
- 复制表, 创建新的 59
- 复制过程
 - 编码标准 58
 - 捕获和应用 57
 - 登台表 58
 - 定制任务 58
 - 注册 62
- 复制和抽取 29
- 复制控制表 61
- 复制源表 60

[G]

- 更改定制 DMS 的缓冲池、表和表空间 19
- 更改缺省报表货币 12
- 更改数据表, 确定大小 59
- 更新 sql 语句 38
- 工具
 - IBM DB2 Intelligent Miner for Data 65
- 购物者 8
- 光标方式, 运行 39

[H]

- 缓冲池, 为 DMS 更改 19

[J]

- 记录, 删除限制 20
- 兼容 DB2 脚本文件的准则 37
- 脚本文件 52, 54
- 进行财政年度修改 19
- 竞销管理 8

[K]

- 控制表
 - APPTABS 69
 - DATATABS 68
 - MODELS 68
- 扩展 43
- 扩展和修订, 定义 43

[L]

- 列 79
 - ABANDONED_MINUTES 71
 - ABANDONED_ORD_STATUS 71
 - ANALYSIS_CATALOG_ID 72
 - BROKER_STORE_ID 72
 - CALENDAR_DATE 17
 - CONTRACT_ACTIVE 72
 - CONTRACT_CANCELLED 72
 - CONTRACT_IN_PREPARATION 72
 - CPN_REDEEMED_ORDER_STATUS 72
 - DAY_OF_FM 17
 - DAY_OF_FM_ID 17
 - DAY_OF_FM_REF 18
 - DAY_OF_FY 18
 - DAY_OF_FY_REF 18
 - DAY_OF_WK 17
 - DAY_OF_WK_ID 17
 - DAY_OF_WK_REF 18
 - DMT_PROSPECT_ORD_STATUS 72
 - DMT_PURCHASER_ORD_STATUS 73
 - EXTRACTION_COUNT 73
 - EXTRACTION_SUCCEED 73
 - FE_EFFECTIVE_MINUTES 73
 - FE_EFFECTIVE_ORD_STATUS 73
 - FISCAL_YR 18
 - FM_EFFECTIVE_MINUTES 74
 - FM_EFFECTIVE_ORD_STATUS 74
 - HOLIDAY_FLG 18
 - MEMBER_ADDRESS_TYPE_ID 74
 - MINING_APPLY_INTERVAL 74
 - MINING_APPLY_TIME 74
 - MINING_BASE_NAME 74

列 (续)

MINING_PASSWORD 74
MINING_TRACE 75
MINING_TRAINING_INTERVAL 75
MINING_TRAINING_TIME 75
MINING_USER_NAME 75
MON_OF_FY 18
MON_OF_FY_ID 18
MON_OF_FY_REF 18
NF_EFFECTIVE_MINUTES 75
NON_PURGE_ORD_STATUS 75
ORDERS_AWAITING_PAYMENT 77
ORDER_STATUS_BILLED 76
ORDER_STATUS_CANCELLED 76
ORDER_STATUS_COLLECTED 76
ORDER_STATUS_ID_NOREV 76
ORDER_STATUS_ID_SUM_MEMBER 76
ORDER_STATUS_ID_SUM_TRADING 76
ORDER_STATUS_XFERRED 77
ORG_BUSINESS_TYPE 77
ORG_INDUSTRY_TYPE 77
PARAM_TYPE 71
PARAM_VALUE 71
PER_AGGR_ID 17
PER_DESC_ID 17
PER_ID 17
PRODUCT_PRICE_AGGREGATE 77
QTR_OF_FY 18
QTR_OF_FY_ID 18
QTR_OF_FY_REF 18
REPLICATION_METHOD 78
REPLICATION_START 78
REPORT_CURRENCY 78
REPORT_LANGUAGE 78
REQUEST_ETL_PROCESS 78
RFM_BINS 78
RFM_INTERVAL 78
RFM_LAST_RUN 78
RFQ_WINNING_RESPONSES 79
ROW_NUM 71
SRF_EFFECTIVE_MINUTES 79
STORE 79
TIME_CUT_OFF 79
TIME_CUT_OFF_LOCAL 80
TIME_CUT_OFF_PREV 80
WCS_SOURCE 80
WEEKDAY_FLG 18
WK_OF_FM 18
WK_OF_FM_ID 18
WK_OF_FM_REF 18
WK_OF_FQ 18
WK_OF_FQ_ID 18
WK_OF_FQ_REF 18
WK_OF_FY 18
WK_OF_FY_ID 18

流量, Web 站点 8

[M]

命令

REORG 23
REORGCHK 23
RUNSTATS 23
模式 (WCA) 66
模型, 注册 67

[P]

配置工具

配置后的参数 81
批处理文件 51, 54

[Q]

启动“捕获”过程 62
前缀, UX 52
趋势 8
确定商店的语言和货币属性 12
群集 66

[R]

日志, 增加 59

[S]

删除记录 20
删除/除去 11
商店的语言和货币属性, 确定 12
商家到商家的交易 7
设置笔记本 67
设置对象 67
使用 ETL 驱动程序 33
市场营销竞选 8
示例
用于群集的数据库模型 66
输入的参数 71
书, WebSphere Commerce vi
数据表, 注册 66
数据集市 13
表的定制约定 49
重新组织表 23
除去 11
描述 3
数据挖掘 65
业务问题 7
在抽取后 71
数据集市定制任务 49
数据集市 (WCA) 66
数据库
备份 11
数据块大小, 更新 23

数据挖掘

创建新模型 67
定制 65
功能 65
新模型, 示例 66
新模型, 注册 68
引进 65
索引类型
主键 68, 69

[T]

特征, 购物者 8
提高性能 23
添加
表 50
列 50
数据 50
突出显示的约定 v

[W]

外壳程序脚本 52
唯一标识 50
文档, WCA v
文件
reference.properties 19
Reference_Table.properties 16
wca_default_period.csv 16
wca_fill_period.sql 16
问题, 业务 7

[X]

系统管理员, 描述 3
销售评估 7, 8
性能, 提高 23
修订 44
修改复制控制表 61

[Y]

业务
报表 7
问题 7
业务报表
访问 3
业务分析师, 描述 3
业务经理, 描述 4
业务问题
产品顾问 8
订单 8
购物者 8
竞销管理 8
类别 7

业务问题 (续)
 商家到商家 7
 销售 8
 Web 站点流量 8
已赋值的 COMPLETED_STATUS 参数
 71
隐喻 8
应用 57
应用结果. 注册 67
用户扩展表 49
用户类型, WCA 3
预测模型标记语言 (PMML) 69
运行复制步骤 63

[Z]

在光标方式下运行 39
在 export / import / load 方式下运行 39
增加日志空间 59
支持 Web 站点 vi
直接映射 14
主键 69
注册
 新数据挖掘模型 68
 APPTABS 表中的应用结果 67
 DATATABS 表中的数据表 66
 MODELS 表中的模型 67
注册复制过程 62
状态列名 69
组合 SQL 语句的方式 39
最近购买时间、购买频率、购买金额
 请看 RFM

[特别字符]

“已完成” 参数值 71
“已装运” 参数值 71

A

ABANDONED_MINUTES 列 71
ABANDONED_ORD_STATUS 列 71
ANALYSIS_CATALOG_ID 列 72
APPTABS 控制表 67, 69
app_type 列名 68

B

BROKER_STORE_ID 列 72

C

CALENDAR_DATE 列 17
CALENDAR_DATE 字段 16

CD 表, 确定大小 59
CommitRows, 定义 35
CommitStatement, 定义 34
CONTRACT_ACTIVE 列 72
CONTRACT_CANCELLED 列 72
CONTRACT_IN_PREPARATION 列 72
CPN_REDEEMED_ORDER_STATUS 列
 72
CURCONVERT, 定义 12
CURLIST, 定义 12
CursorInsert 39
CursorUpdate 39

D

DATATABS 控制表 66, 68
data_desc 列名 68
data_id 列名 68, 69
data_location 列名 68
data_name 列名 68
data_type 列名 68
DAY_OF_FM 列 17
DAY_OF_FM_ID 列 17
DAY_OF_FM_REF 列 18
DAY_OF_FY 列 18
DAY_OF_FY_REF 列 18
DAY_OF_WK 列 17
DAY_OF_WK_ID 列 17
DAY_OF_WK_REF 列 18
DB2
 仓库中心脚本 52, 54
 控制中心 11
 Intelligent Miner for Data 65, 66
DB2 通用数据库
 支持 Web 站点 vi
DMT_PROSPECT_ORD_STATUS 列 72
DMT_PURCHASER_ORD_STATUS 列 73

E

ETL
 抽取、转换、移动和装入过程 52
 过程, 描述 25
 流 27, 29
 修改过程 44
ETL 驱动程序
 错误处理 40
 使用 33
 受支持的本地命令 34
 受支持的命令标记 33
 受支持的全局命令 34
export / import / load 方式, 运行 39
EXTRACTION_COUNT 列 73
EXTRACTION_SUCCEED 列 73

F

FE_EFFECTIVE_MINUTES 列 73
FE_EFFECTIVE_ORD_STATUS 列 73
file_location 列名 69
FISCAL_YR 列 18
FM_EFFECTIVE_MINUTES 列 74
FM_EFFECTIVE_ORD_STATUS 列 74

H

HOLIDAY_FLG 列 18

I

Intelligent Miner for Data
 数据挖掘功能 3
 与 WCA 一起工作 5
 支持 Web 站点, vi

L

LAST_UPDATED_REF 表 50
LoadInsert 40
LoadInsert, 定义 37
LoadOption, 定义 36
LoadParam, 定义 36
LoadUpdate 40

M

MAX_SYNCH_MINUTES, 更新 23
MEMBER_ADDRESS_TYPE_ID 列 74
MINING_APPLY_INTERVAL 列 74
MINING_APPLY_TIME 列 74
MINING_BASE_NAME 列 74
MINING_PASSWORD 列 74
MINING_TRACE 列 75
MINING_TRAINING_INTERVAL 列 75
MINING_TRAINING_TIME 列 75
MINING_USER_NAME 列 75
MNF_EFFECTIVE_ORDER_STATUS_ID
 列 75
MODELS 控制表 67, 68
model_alg 列名 69
model_desc 列名 69
model_id 列名 69
model_name 列名 69
model_type 列名 69
MON_OF_FY 列 18
MON_OF_FY_ID 列 18
MON_OF_FY_REF 列 18

N

NON_PURGE_ORD_STATUS 列 75

O

ORDERS_AWAITING_PAYMENT 列 77
ORDER_STATUS_BILLED 列 76
ORDER_STATUS_CANCELLED 列 76
ORDER_STATUS_COLLECTED 列 76
ORDER_STATUS_ID_NOREV 列 76
ORDER_STATUS_ID_SUM_MEMBER 列 76
ORDER_STATUS_ID_SUM_TRADING 列 76
ORDER_STATUS_XFERRED 列 77
ORG_BUSINESS_TYPE 列 77
ORG_INDUSTRY_TYPE 列 77

P

PARAMETER 表 72
ABANDONED_MINUTES 列 71
ABANDONED_ORD_STATUS 列 71
ANALYSIS_CATALOG_ID 列 72
BROKER_STORE_ID 列 72
CONTRACT_ACTIVE 列 72
CONTRACT_CANCELLED 列 72
CONTRACT_IN_PREPARATION 列 72
DMT_PROSPECT_ORD_STATUS 列 72
DMT_PURCHASER_ORD_STATUS 列 73
EXTRACTION_COUNT 列 73
EXTRACTION_SUCCEED 列 73
FE_EFFECTIVE_MINUTES 列 73
FE_EFFECTIVE_ORD_STATUS 列 73
FM_EFFECTIVE_MINUTES 列 74
FM_EFFECTIVE_ORD_STATUS 列 74
MEMBER_ADDRESS_TYPE_ID 列 74
MINING_APPLY_INTERVAL 列 74
MINING_APPLY_TIME 列 74
MINING_BASE_NAME 列 74
MINING_PASSWORD 列 74
MINING_TRACE 列 75
MINING_TRAINING_INTERVAL 列 75
MINING_TRAINING_TIME 列 75
MINING_USER_NAME 列 75
NF_EFFECTIVE_MINUTES 列 75
NF_EFFECTIVE_ORDER_STATUS_ID 列 75
NON_PURGE_ORD_STATUS 列 75

PARAMETER 表 (续)

ORDERS_AWAITING_PAYMENT 列 77
ORDER_STATUS_BILLED 列 76
ORDER_STATUS_CANCELLED 列 76
ORDER_STATUS_COLLECTED 列 76
ORDER_STATUS_ID_NOREV 列 76
ORDER_STATUS_ID_SUM_MEMBER 列 76
ORDER_STATUS_ID_SUM_TRADING 列 76
ORDER_STATUS_XFERRED 列 77
ORG_BUSINESS_TYPE 列 77
ORG_INDUSTRY_TYPE 列 77
PARAM_TYPE 列 71
PARAM_VALUE 列 71
PRODUCT_PRICE_AGGREGATE 列 77
REPLICATION_METHOD 列 78
REPLICATION_START 列 78
REPORT_CURRENCY 列 78
REPORT_LANGUAGE 列 78
REQUEST_ETL_PROCESS 列 78
RFM_BINS 列 78
RFM_INTERVAL 列 78
RFM_LAST_RUN 列 78
RFQRSP_OUTSTANDING_ORDERS 列 79
RFQ_RESPONSE_IN_PREPARATION 列 79
RFQ_WINNING_RESPONSES 列 79
ROW_NUM 列 71
SRF_EFFECTIVE_MINUTES 列 79
SRF_EFFECTIVE_ORDER_STATUS_IDS 列 79
STORE 列 79
TIME_CUT_OFF 列 79
TIME_CUT_OFF_LOCAL 列 80
TIME_CUT_OFF_PREV 列 80
WCS_SOURCE 列 80
PARAM_TYPE 列 71
PARAM_VALUE 列 71
PER_AGGR_ID 列 17
PER_DESC_ID 列 17
PER_ID 列 17
per_id 列名 69
PER_ID 字段 16
Primary Keys, 定义 36
PRODUCT_PRICE_AGGREGATE 列 77

Q

QTR_OF_FY 列 18
QTR_OF_FY_ID 列 18
QTR_OF_FY_REF 列 18

R

reference.properties 文件 19
Reference_Table.properties 文件 16
REORG 命令 23
REORGCHK 命令 23
REPLICATION_METHOD 列 78
REPLICATION_START 列 78
REPORT_CURRENCY 列 78
REPORT_LANGUAGE 列 78
rep.sql 文件 59
REQUEST_ETL_PROCESS 列 78
res_file_name 列名 69
RFM
定义 65
确定值的条件 70
RFM_BINS 列 78
RFM_INTERVAL 列 78
RFM_LAST_RUN 列 78
RFQRSP_OUTSTANDING_ORDERS 列 79
RFQ_RESPONSE_IN_PREPARATION 列 79
RFQ_WINNING_RESPONSES 列 79
ROW_NUM 列 71
RUNSTATS 命令 23
Runstats, 定义 34

S

sample_pct 列名 68
SelectPrimaryKeys, 定义 37
SourceKeys, 定义 35
SourceTable, 定义 35
SRF_EFFECTIVE_MINUTES 列 79
SRF_EFFECTIVE_ORDER_STATUS_ID 列 79
STORE 列 79
STORE, 定义 12
STOREENT, 定义 12
STORELANG, 定义 12

T

table_id 列名 69
table_name 列名 69
table_type 列名 69
TargetKeys, 定义 36
TargetTable, 定义 36
TIME_CUT_OFF 列 79
TIME_CUT_OFF_LOCAL 列 80
TIME_CUT_OFF_PREV 列 80

U

%IWDA_DIR%, 定义 v

UpdateCondition, 定义 37
UpdateSource, 定义 35
UpdateType, 定义 37
usage_type 列名 68
UX 前缀 52
UX 用户扩展表 50
UX 用户扩展列 50

W

WCA

备份数据库 11
重新组织表 23
概述 3
模式 66
数据集市 3, 66
随同提供的软件 3
文档 v
业务报表 3
用户类型 3
与 WebSphere Commerce 一起工作 4

WCA 表
 请看 表

WCA 服务器
 提高性能 23

WCA 列
 请看 列

WCA ETL 源 29

WCA .PARAMETER 表 81

WCA.PARAMETERS 表
 列名 71

wca_default_period.csv 文件 16

wca_fill_period.sql 文件 16

WCS_SOURCE 列 80

Web 站点流量 8

Web 站点, 支持 vi

WebSphere 贸易加速器 3

WebSphere Commerce
 事务数据库服务器 4
 书 vi
 网上商店 3
 与 WCA 一起工作 4
 WebSphere 贸易加速器 3

WEEKDAY_FLG 列 18

WK_OF_FM 列 18

WK_OF_FM_ID 列 18

WK_OF_FM_REF 列 18

WK_OF_FQ 列 18

WK_OF_FQ_ID 列 18

WK_OF_FQ_REF 列 18

WK_OF_FY 列 18

WK_OF_FY_ID 列 18

.bat (批处理) 文件 51, 54

.sh (外壳程序) 脚本 52



中国印刷