**目 录**

[**1.** 引言 1](#_Toc489122319)

[1.1 实习目的和要求 1](#_Toc489122320)

[1.2 实习内容 1](#_Toc489122321)

[1.3 实习单位概况（在外单位实习的请描述相关单位概况） 3](#_Toc489122322)

[**2.** 专题讲座心得 5](#_Toc489122323)

[2.1 “大数据技术：问题，技术及挑战”专题讲座 5](#_Toc489122324)

[2.2 “网络空间安全”专题讲座 6](#_Toc489122325)

[2.3 “下一代互联网：机遇与挑战”专题讲座 6](#_Toc489122326)

[**3.** 北京天河自行车销售模块数据分析 9](#_Toc489122327)

[3.1 项目概述 9](#_Toc489122328)

[3.1.1 需求分析 9](#_Toc489122329)

[3.1.2 数据逻辑模型 9](#_Toc489122330)

[3.2 数据源分析 10](#_Toc489122331)

[3.3 数据仓库设计 12](#_Toc489122332)

[3.4 ETL开发与设计 12](#_Toc489122333)

[3.5 报表的开发与设计 16](#_Toc489122334)

[3.6 报表展示 18](#_Toc489122335)

[**4.** 实习总结 26](#_Toc489122336)

[4.1 实习内容的复杂性评价 26](#_Toc489122337)

[4.2 实习体会、收获与建议 26](#_Toc489122338)

[**5.** 参考文献 30](#_Toc489122339)

[**6.** 附录 实习日记 31](#_Toc489122340)

# 引言

## 实习目的和要求

1. 实习目的：

计算机科学与技术是一门理论和实践相结合的学科，经过了大学三年的沉淀之后，我们掌握了相当多的理论知识，但是一直缺少一个真正应用它们的平台，于是在大三的结束学校为我们精心准备了生产实习这门实践性很强的科目，让我们对自己的学习能力和团队协作能力进一步升华。

生产实习是本科教学计划中非常重要的实践性教学环节，是专业培养目标和教学计划、课程设置的有机组成部分，是理论教学的完善和补充。通过生产实习，使学生通过实践了解和掌握计算机的应用的开发、设计、运行等环节。通过生产实习培养学生掌握从事计算机系统开发的技能，加深对已经学习过的专业理论知识的理解和认识，提高在生产实践中调查研究、观察问题分析问题以及解决问题的能力和方法。为学生后续学习打下基础。

(2) 实习要求：

邀请专家做专题报告，了解计算机专业各领域发展方向和前景，主要研究内容和应掌握的理论与技术。

参观东软集团，了解东软主要发展历程，企业发展方向。学习他们的创业、敬业精神。

运用上课所学的sql知识和实习期间学习的有关数据仓库的理论，建立一个机遇ETL工具的商务智能数据分析模块。站在数据库管理者的角度，要求实现对报表的功能性，复杂性和美观程度的兼容，最后做到能够像没有接触过数据库的人清晰地展示数据的统计情况的功能。

## 实习内容

学习sql语句的使用，使用PLSQL联系对数据的增删改查，搭建数据仓库，进行维度建模（星型结构），使用informatica（ETL工具）对数据项进行抽取，转换，加载，在biee上完成报表的建立，使用各种图形描述数据信息，最后建立仪表盘对数据进行整合。

## 实习单位概况

东北大学计算机学院实验教学中心简介：东北大学计算机学院实验教学中心承担着计算机学院计算机类专业和电子信息类专业的实验教学任务。此外，还承担各类课程设计、生产实习和毕业实习等实践教学任务。2015年东北大学计算机实验教学中心入选国家级实验教学示范中心。

东北大学计算机科学与工程学院于2015年12月成立，其历史可追溯至1923年建校时设立的电工学系，并于20世纪50年代创建了无线电专业和计算机专业。在计算机科学与技术、信息与通信工程两个一级学科的发展历程中，创造了一个又一个值得载入史册的辉煌成绩：成功研制了我国第一台模拟电子计算机，自主研制了我国第一台CT机，在国内第一批设立无线电专业，培养出我国第一批计算机专业本科毕业生，设立国家第一批计算机应用博士点并培养出我国第一位计算机应用博士，孕育了国内第一家软件上市公司，建立了中国第一个局部互联网……

学院现有教职工168人，其中专任教师118人，全日制学生3498人。设有计算机科学系、计算机工程系、通信与电子工程系3个系，大数据及人工智能技术、嵌入式系统及安全互联、医学信息系统及数字医疗和教学研究中心等4个研究中心。学院有图灵奖获得者（法国科学院院士，法国工程院院士，欧洲科学院院士，美国科学院院士，IEEE Fellow）1人、欧洲科学院院士（IEEE Fellow）1人、国务院学科评议组成员1人、国家“千人计划”入选者2人、“长江学者奖励计划”特聘教授2人、国家杰出青年科学基金获得者2人、国家百千万人才工程百人层次2人、万人计划青年拔尖人才1人、国家优秀青年基金获得者2人、教育部跨（新）世纪优秀人才13人、辽宁省教学名师2人。此外，还有包括2名院士在内的几十位国内外著名学者在学院兼职任教。

学院现有计算机科学与技术、通信工程、电子信息工程、物联网工程4个本科专业，有计算机科学与技术、信息与通信工程两个一级学科博士学位授权点和博士后工作流动站。计算机应用技术为二级学科国家重点学科，计算机科学与技术为一级学科辽宁省重点学科，通信与信息系统为二级学科辽宁省重点学科。计算机学科拥有和覆盖“软件架构新技术国家重点实验室”“计算机软件国家工程技术研究中心”“数字化医学影像设备国家工程技术研究中心”“国家级计算机实验教学示范中心”等13个国家、省部级学科和教学实验基地。计算机科学与技术一级学科在三次全国一级学科整体水平评估中两次位居第九、一次位居第十。2013年、2015年，计算机科学与技术学科两次进入ESI排名全球前1%。

近年来，学院科研氛围日益浓厚、科研规模日益扩大、科研水平日益提升，累计承担国家级科研项目175项，科研经费近5亿元，获得国家科技进步二等奖4项。

学院积极开展多渠道、多层次的国际交流与合作，邀请大批国际知名专家来校开展交流合作。教师出国讲学、合作科研和进修的人数不断增加，学生作为交流生、互派生到国际知名院校深造的机会不断增多。

学院注重学生创新创业氛围营造，学生在国际大学生程序设计竞赛、美国大学生数学建模竞赛、“挑战杯”等国际国内大赛中获奖近千人次。

面向未来，东北大学计算机科学与工程学院将以“产学研、国际化”为特色，以一流的队伍、一流的教学、一流的科研和一流的管理体制和运行机制为保障，全面提升人才培养质量，把握发展机遇，拓宽学科方向，扩大国际影响，把东北大学计算机建成“国内一流、国际知名”的计算机学科。

# 专题讲座心得

在生产实习期间，我们聆听了多位专家学者以及企业主管的讲座和报告。对此，感谢学校让我们有机会现场聆听各位专家学者对于计算机专业各领域的研究成果和个人见解，同时给我们提供了一个学术交流的平台，这些演讲将我们平时的学习成果和现实中的应用结合在一起，使我们受益颇多。

首先，是于戈教授的大数据分析：问题、技术及挑战专题讲座；接下来是姚羽教授对当代互联网有关发展的讲座；然后是王兴伟教授的下一代互联网：机遇与挑战专题讲座。以下是我听取这些讲座的心得：

## 大数据技术专题讲座

作为我们专题讲座的第一讲，计算机学院的院长于戈教授无疑为我们带来了一场精彩的讲座，不论是从对计算机科学的理解还是演讲的深度广度，于戈教授倾其所有的为我们展示一个新时代背景下的多喝嵌入式社会会具有的特征，在未来，我们将担负起来人类社会进步的重要使命，我们也一定会.不辱使命，在为欸来的社会发展中争取贡献自己的一份力量。

“[大数据](http://baike.baidu.com/view/6954399.htm)”作为时下最火热的IT行业的词汇，随之而来的数据仓库、数据安全、数据分析、数据挖掘等等围绕大数据的商业价值的利用逐渐成为行业人士争相追捧的利润焦点。

。于老师指出，在大数据时代，每一个个人都由许许多多的数据所围绕，而这些数据的大小，数量都是我们人类所难以想象的，我们的健康，信用，财产，性格，工作，家庭，都要靠这些数据连接在一起，而今后这些代表了我们的资产的数据的管理也就显得尤为关键，试想，如果黑客删除了我们的银行账户的存款，那么我们的资产就收到了损害，虽然大数据为我们的生活带来了进步，但是伴随着进步我们也将面对像上面姚羽教授所提到的那样的问题，而这些问题正是我们这些新时代的计算机学生要去解决的。

未来就在我们的手中，这是最好的时代，也是最坏的时代。这句话是计算机网络结课的时候姚羽老师送给我们的，在当天的讲座中，姚羽老师也语重心长的鼓励我们在大学期间重点培养自己的兴趣，如果能够发觉自己对计算机网络的热爱，然后投身其中，那必然会让我们在工作中有更大的动力和干劲儿。

而这就是一种传承的力量，在任何一个学科的发展的道路上都需要我们前赴后继的为我们共同的伟大理想做出贡献，于戈老师将火炬传到了我们的手中，我们要通过自己的实习行动，将东北大学的旗帜传承下去。

## 网络空间安全专题讲座

网络空间安全技术的讲座是由我们的计算机网络老师姚羽带来的，做为一位非常受欢迎的老师姚羽老师还是我们大数据研究所的主任，平时总是非常幽默的姚羽老师这次也是以一种风趣的形式为我们带来这次演讲。

在未来，网络安全会成为我们生活中至关重要的组成部分，我们生活中的很多可以用数据描述的数字存放在电脑和网络中，我们需要一个安全的网络环境来存储，操作我们的数据。而现在我们的网络环境却没有我们最终期待的那样的安全，黑客现象还是时有发生，网络诈骗还是风头正劲，姚羽老师作为我们的计算机网络技术的老师，风趣幽默的为我们解释了什么是网络安全和网络安全在我们生活中所体现出来的重要性。

如果从事这个方向的话，姚羽老师指出，网络空间安全技术是在不断更新和变化，我们所要面对的不单单是想要搞破坏的黑客们，还有日益增长的数据量带来的存储，运算上的复杂度的增加，我们的的目标是维持一个安全干净的网络环境，所以这就需要我们在学习的过程中考虑到任何可能遇到的问题。

我们的姚羽教授最后还引用了金庸小说中的人物来描述网络安全的层次：你知道你知道，你知道你不知道，你不知道你不知道。我想，这不但但对于我们网络安全，对于我们在学习中也是这个道理，我们只有明确了我们想要达成的目标，并且在实干的道路上并不断完善自己的思维，总计问题，最后才能充分的了解我们想到得到的结果，缺德我们预期的成绩。我想这也是姚羽教授想要教会我们的道理。

## “下一代互联网：机遇与挑战”专题讲座

互联网是我们日常生活中不可缺少的一部分。随着互联网的普及，人们对互联网的以来也越来越严重。

此次演讲的王兴伟教授是一个风趣幽默的教授，整个讲座的过程中，他都在和学生们进行着良好的互动，在整个演讲过程中气氛轻松愉快，不过王教授也同时提到了一些严肃的问题要我们去思考，的确，下一代互联网：机遇与挑战并存，我们要抓住机遇，面对挑战，在复杂的环境中找准方向。

现在的信息行业飞速发展，在过不多久就将进入5G时代，而像5G技术一样的更多的新鲜技术也将进入我们的生活，让我们感受到计算机带给我们的切实的方便。

互联网时代的到来，新科学技术革命引起的新产业革命同步发展的机遇，豆浆在今后我们步入社会，走上岗位的时候伴随我们左右，不论我们是选择读研究生还是工作，我们都在以自己的方式为我们的计算机科学做自己的贡献，我们要时时刻刻都保存这一种求实务实的态度，这样我们才能在我们的工作中取得我们理想的成功。

王兴伟教授在这次演讲中向我们介绍了如今的互联网发展形势：我们普遍使用的IPV4正在慢慢消耗殆尽，IPV6做为新兴的标准真正在慢慢兴起，IPV6继承了IPV4的优点并且将要摒弃IPV4在网段划分上的一些问题，使之成为更适合新需要的网络协议。

通过这次专题讲座，我对互联网有了更深入的了解。我会更加学好专业知识，拓展自己的知识面，让自己能更好地适应时代的潮流。

结束了三次生动的汇报，我感觉到了我们作为新时代的计算机学子在肩上使命的重大，作为一名东北大学的计算机学生，我想在今后的道路上多发样我们的自强不补习，知行合一的精神，为我们的母校争光添彩的同时，也争取能够实现我为祖国做出自己的贡献的目标。

# 北京天河自行车销售模块数据分析

## 项目概述

### 背景

**公司背景：**

北京天河自行车制造有限公司（简称天河自行车），是一家自行车生产制造及销售的企业，目前公司的业务已经覆盖了全国25个省市的超过1000家企业。公司主营业务为自行车的制造及销售，公司的业务流程包括了，材料的采购、铸造、自行车成品及备件的加工、制造等多个环节。在中国自行车制造行业处于领军地位**。**

**项目背景：**

公司自2010年采用ERP系统以来积累了大量采购、生产、销售、以及公司内部运营等数据，随着业务的不断扩充，数据量的增多，使各部门的数据分析也越来越复杂。随着公司业务发展的扩充以及向海外发展的长远目标，公司各阶层领导已经逐步意识到经验促使决策的片面性和不确定性。信息部门结合公司的当前业务运营管理需求与健康扩张（做大的同时，更需要做强）的战略，系统地设计与开发各类关键基础运营（原始）数据的管理分析报表体系与相应的基础运营数据集中管理平台。

### 需求分析

**业务需求：**

天河自行车集团组织架构丰富、业务多样、业绩增长快速。公司决策层面及业务负责人、部门负责人在决策管理方面的需求具有多样化，各部门业务间的需求差异性大，根据决策管理需要，天河自行车数据决策支持平台从各业务、部门等维度出发，将各部门具体业务需求模块化，在实施过程中分模块进行需求调研与分析。

**主要目标**

本期主要进行业务分析决策平台的开发，平台功能需达到使公司的供应商管理、门店绩效管理、各商品品类业绩考核建立在可靠的业务数据基础上，使公司的相关职能部门员工与管理者能够及时看到他们所关心的业务数据，从而为企业决策者正确决策提供准确、及时的依据，并使企业能更好地为供应商和顾客服务，进而提高企业的经营管理水平，提升企业整个供应链的竞争力。

**功能需求**

依据ERP系统中销售模块数据收集，以及公司销售订单业务流程，进行公司贸易整合。在决策系统中达成销售订单管理、监控；产品出库、结算流程监控；并根据销售订单进行采购计划规划，优化[采购]→[入库]→[生产]→[出库]→[销售]整条链路，进而达到减少损耗、节约成本、简化流程作用。商贸系统要求如下：

1）销售订单

主要功能包括接收、更新、保存来自于ERP系统的销售订单数据，为后续销售业务提供

基础数据。系统支持对销售订单的实时跟踪，以更好的把控整个销售流程。

2）产品库存

为了完成产成品出厂发运任务，支持对产成品信息的统一和全程管理。在物料管理模块

的支撑下，收集生产厂厂内成品的产出信息，并且集成厂外库的仓库信息，形成统一的

产成品库信息。成品库存及原料库存信息主要来源于生产管理系统的实时物料库存信息。

3）销售出库

为指导销售出厂，系统接收来自ERP的发货计划（即发货提单），并转化成商贸系统的

物流任务，并由发货单位在物料模块中执行具体的发货任务。

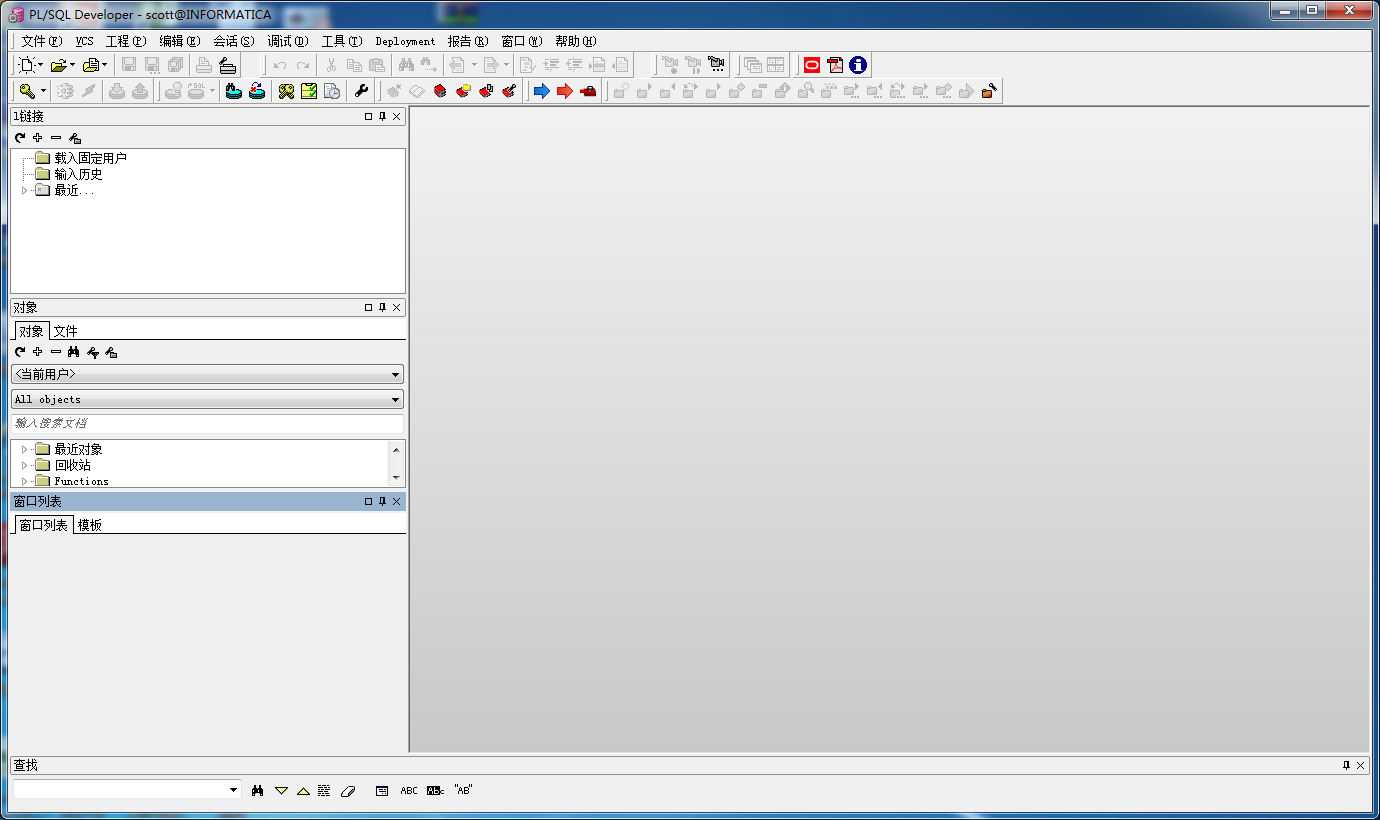
4）销售结算

根据发货出库明细形成销售结算清单，作为销售结算需要结齐的单据，并且可形成出库

明细汇总后的形式发票，上传给ERP系统完成销售结算流程。

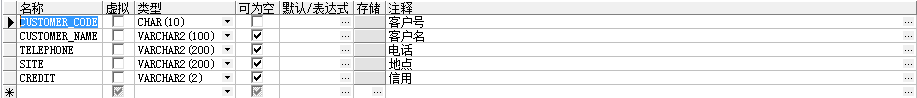
## 数据源设计

首先我们使用PL/SQL熟悉对于sql语句的掌握，强化了一些sql技能

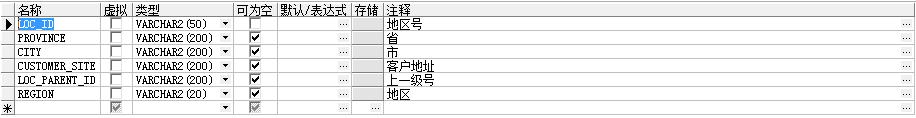


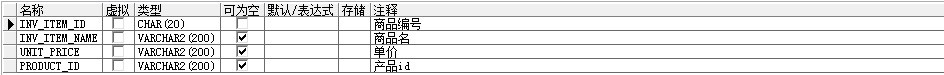
我们建立了一些用于存储数据的目标表：

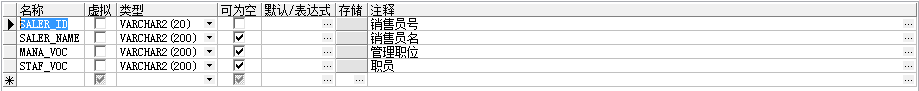
Customer表：



Loc\_table表：

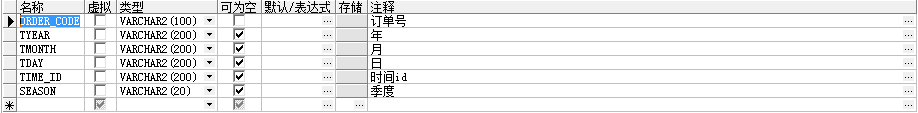


Product表：  


Saler表：  


Sale\_table表：  

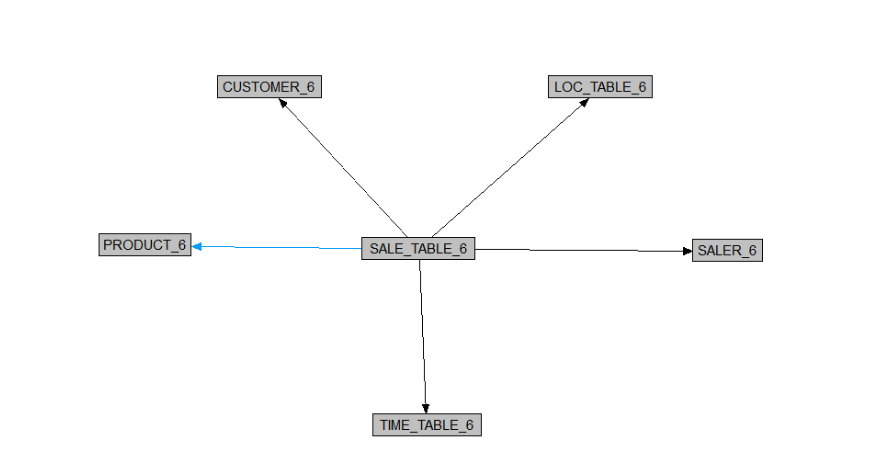

Time\_table表：



## 数据仓库设计

我们根据老师实习时介绍的关系数据库理论，建立了一个一星型模式为基础的数据仓库模型，将我们的Sale\_table表作为事实表，将其与的表作为描述销售信息的五个维度进行分析。

表3.1 维度表建模结构



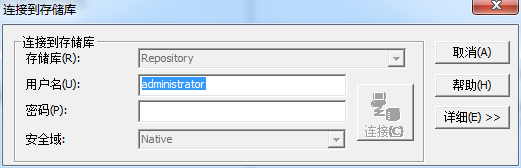
## ETL开发设计

模块设计：主要写各个功能模块是如何设计的，包括实现算法、程序流程图等，每一个模块用到的数据表、类（及类图）、构件（如果有的话），如果某模块比较复杂，画出此模块的程序流程图、时序图等。

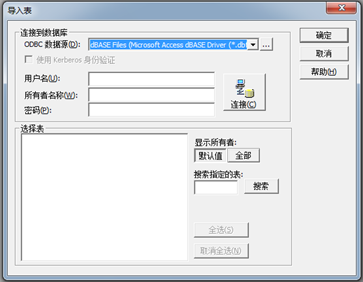
还有界面设计、数据库连接设计等。

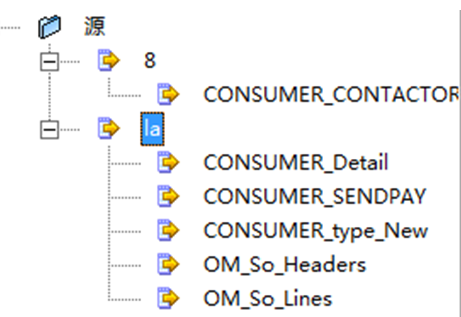
这个部分是整个这次实习的重点部分，我们需要将源数据库中的数据导入实现建好的目标表中，在导入的过程中我们需要注意两张表的对应关系，包括键值约束对应关系，数据类型对应关系，数据对应关系。

首先我们使用ETL工具informatica连接到存储库：

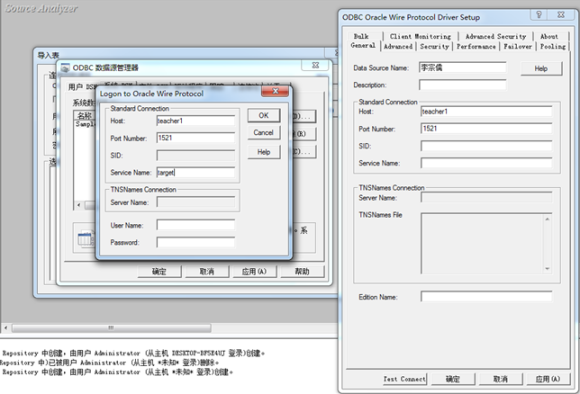


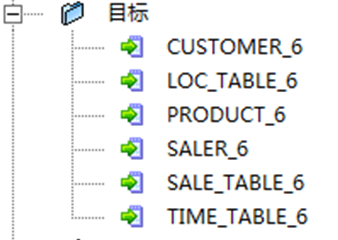
我们将我们的源数据表导入到informatica中：





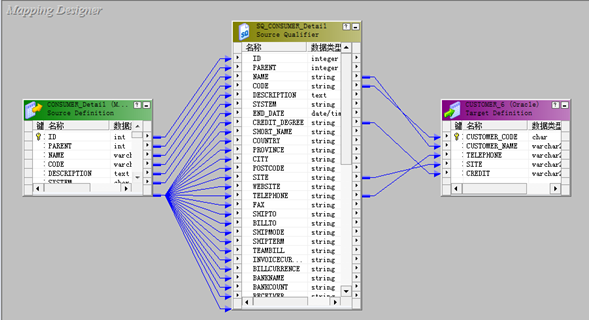
之后我们导入目标表：



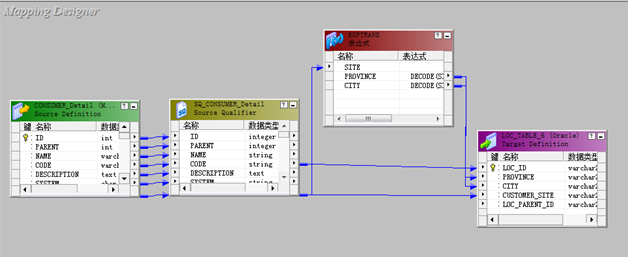


然后我们根据数据之间的关系建立映射：

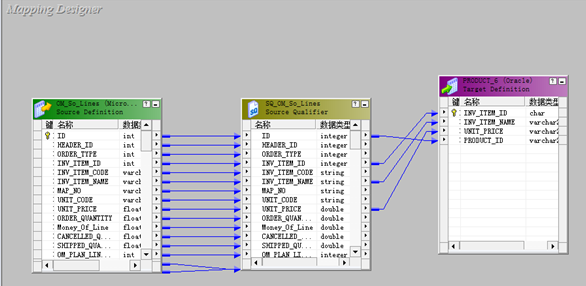
Csutomer表映射：



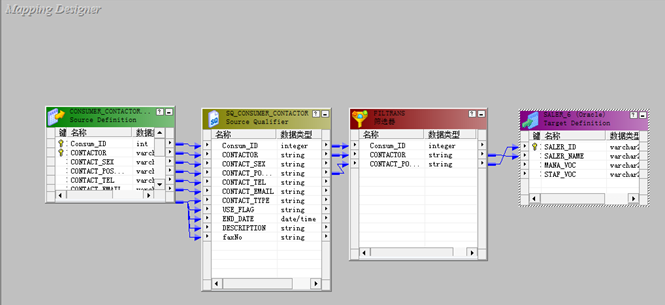
Loc\_table表映射：



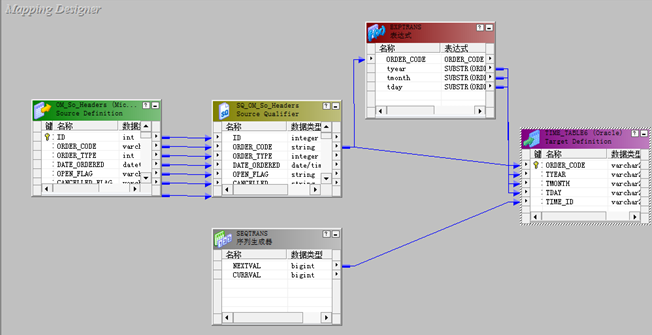
Product表映射：



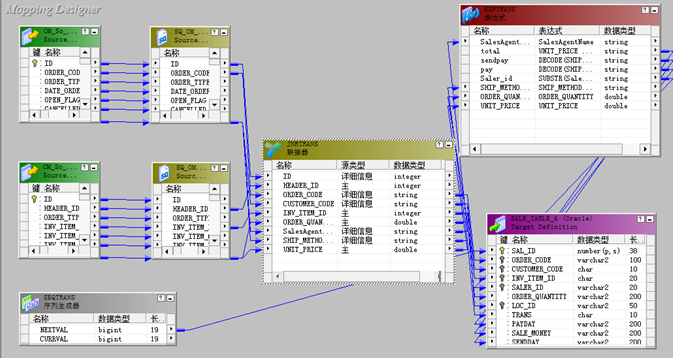
Saler表映射：



Time\_table表映射：

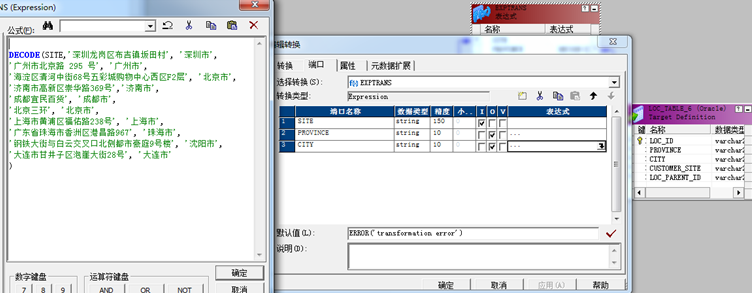
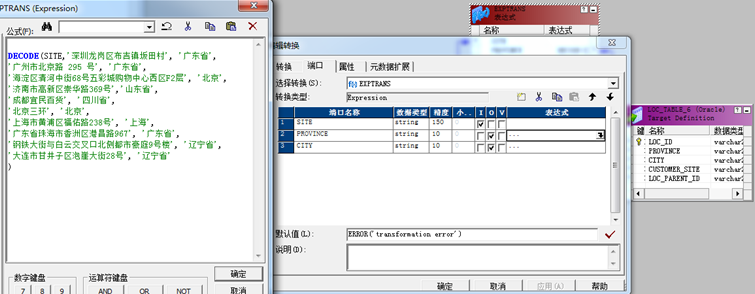


Sale\_table表映射：

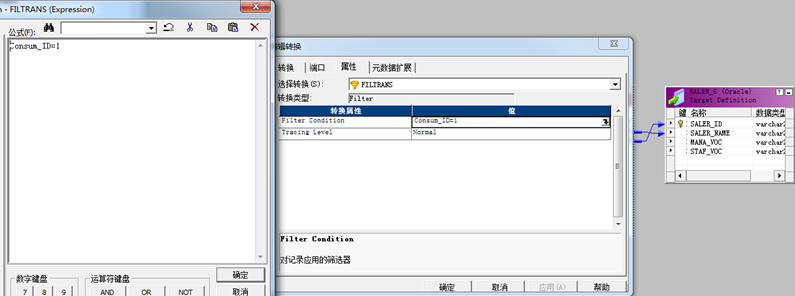


在建立映射的过程中我们需要对信息进行处理，是的他们满足目标表对数据的要求：

1）：在Loc\_table表中，我们用到decode函数将客户的具体地址细分出：大区，省，市，区这些层级，并且设置一个地址ID唯一标识一个用户的地址。



2）：在销售员表saler表中，我们需要使用筛选器将我们选定的销售具体商品的员工筛选出来，在这里，有约我们用到的数据满足销售员的ID号都等于1，所以我们可以直接令consumer ID=1，即可得到销售员信息表：



3）：在时间表time\_table表中，我们也需要划分时间概念上的层级，而这个层级在后续的工作中特别重要，所以我们使用substr函数，将产品销售订单号截取出有效的字符串，之后将它转化成整型（在后续的工作中体现），在这里，我们需要考虑到一个时间能否唯一标识一条订单信息，因为订单信息众多，订单号是精确到分钟之后还有其他值标识来区分的（比如：秒），而我们只是截取到了天这个层级，所以要进行去重操作，保证在后续同比环比的度量值构建中时间维度能正常使用。

截取年：



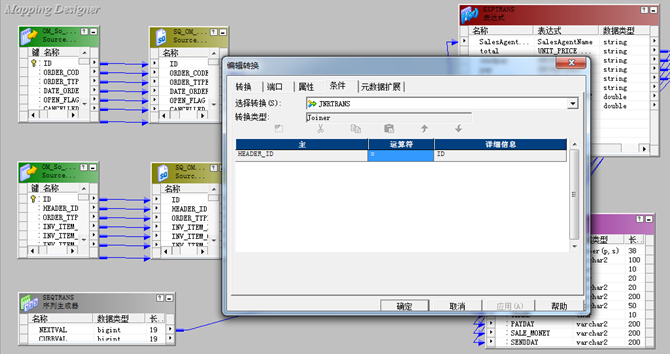
截取月份



截取天：



4）：在事实表sale\_table表中，我们需要使用连接器将两张源数据表中的信息连接在一起，这时候我们需要确定我们链接的属性是什么：（ID&Header\_ID）



事实表中的统计量众多，我们要在expression表达式组件中对他们进行编辑，首先我们需要编辑产品的销售额，我们使用单价乘以销售量的方式：



我们还设计了运费和运输天数这些度量值，我们用decode函数对这个值进行赋值：

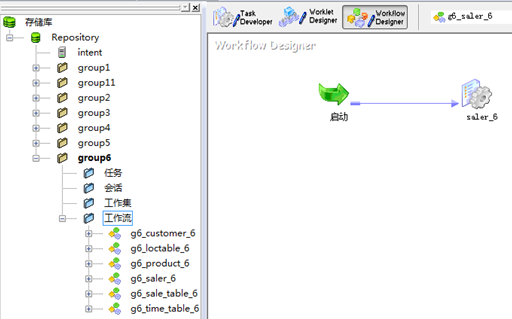


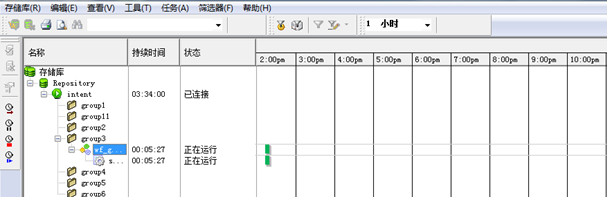


在销售员等级这个值得设计中，由于我们的销售员等级在salesagentname字段中，我们只需要截取两个字符，就能得到它：

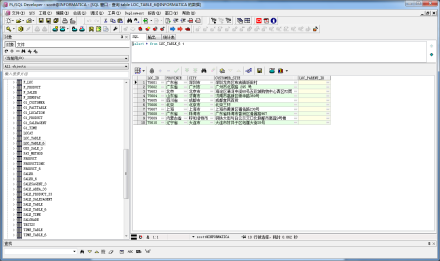
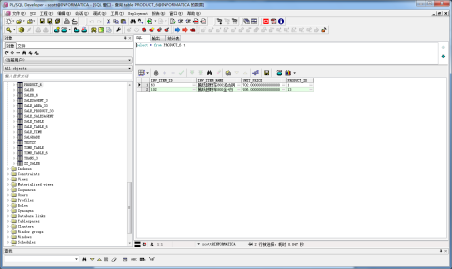
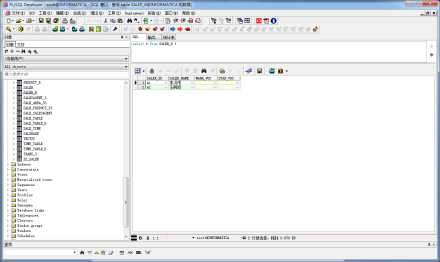
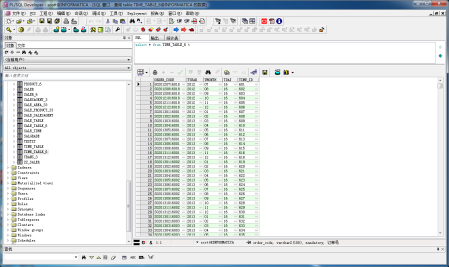
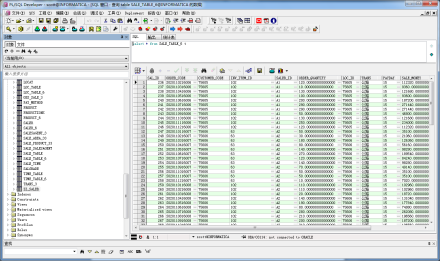
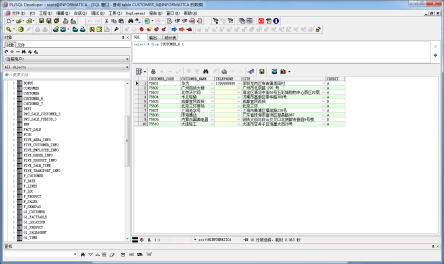


在整理完成之后我们生成工作流，将源数据导入到目标表中：



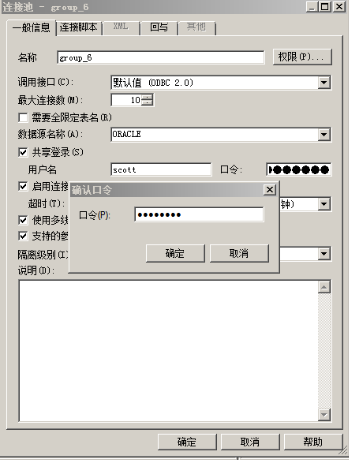


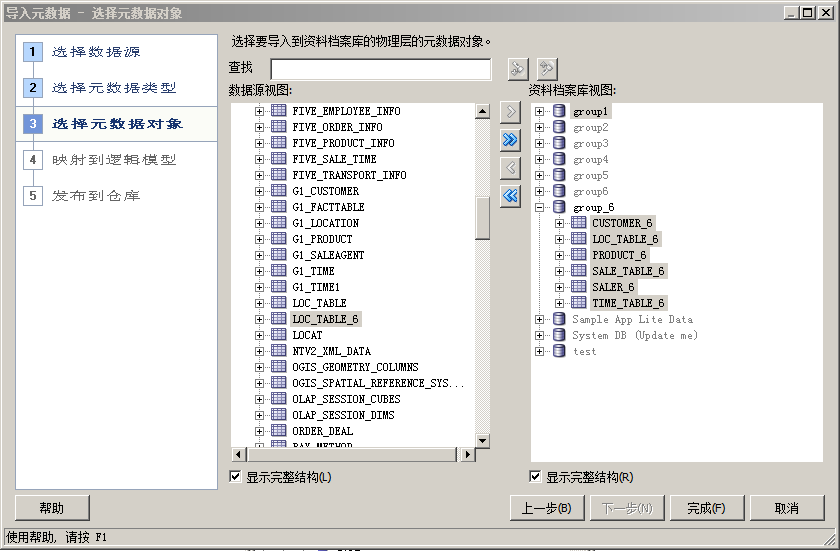
工作流导入成功！接下来在目标表中就会有我们需要的数据：

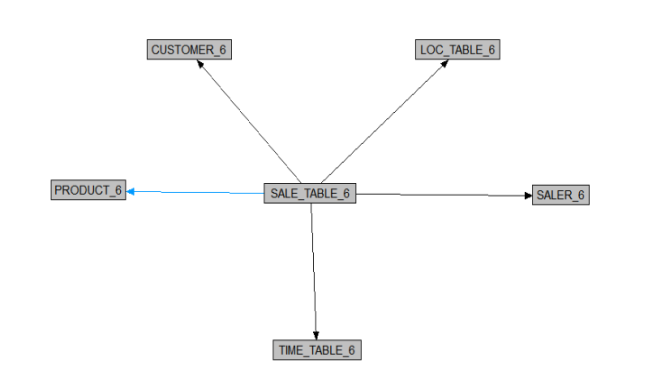
## 报表开发与设计

我们已经有了制作报表的表格和数据，现在我们需要把报表连接到我们的数据库中，用BIEE工具对数据进行分析，首先我们在连接池中将我们的目标表导入：

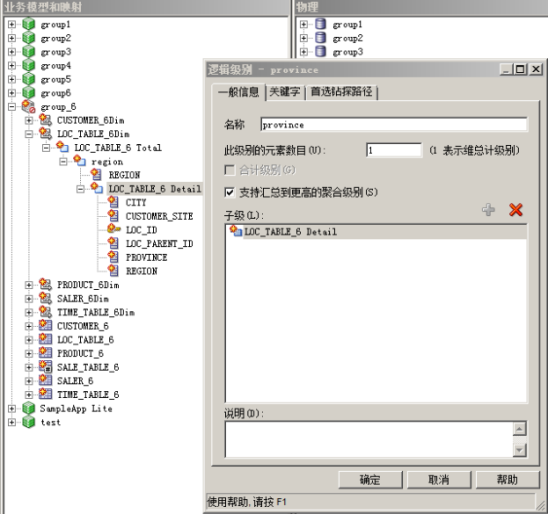




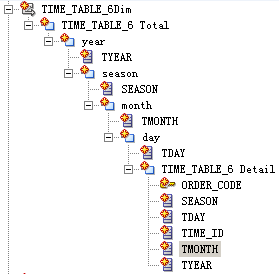
我们按照我们实现设计好的星型模型建模，事实表是SALE\_TABLE\_6，其他表作为维度表与事实表相关联：



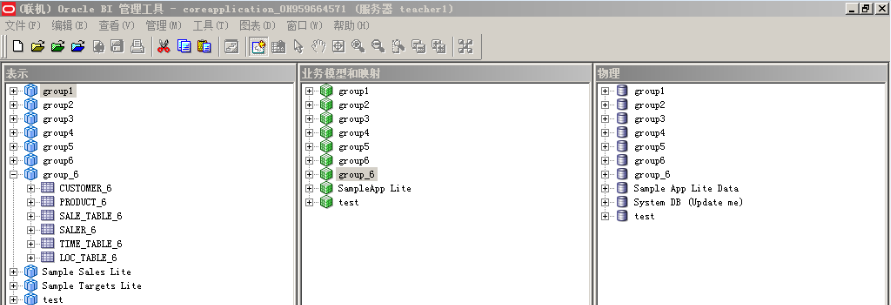
为了能在报表中展示我们的层级关系，我们需要在数据导入阶段对那些有等级关系的数据添加层级约束：



在这里，添加层级约束就是指将更高一层的度量值设置为它的第一层的度量值的父级：



导入之后我们可以在BIEE网页上使用我们的数据：

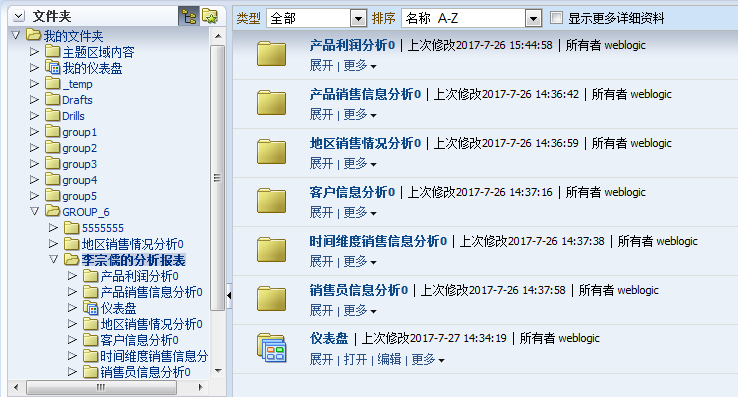


## 报表展示

我们在OBIEE上进行模型建立以及报表开发：



这是一款功能十分强大的软件，我们可以在“我的文件夹”新建自己的分析，在主题区域中可以看到自己的目标表，其中的属性和数值是可以编辑的，还可以根据需要增加或者修改一些属性列，这样不仅可以给一些数据做标记，还可以改正一些在ETL过程中没有注意到的不足，是非常人性化的一个功能。



我将我的分析表按照维度表的分类进行分区，并且加入了一些新的内容，共分成产品利润分析。产品销售信息分析，地区销售情况分析，客户信息分析，时间维度销售信息分析，销售员信息分析，细节表，利润分析八个分析表：

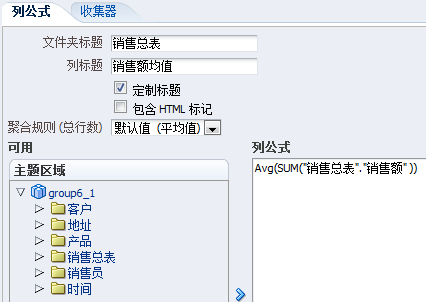
接下来我将对自己的特色进行展示：

1.新增属性列：

1）：平均值列属性：

使用sum和avg函数，求出每个度量值的均值反应在表格中，直观的反映出来单个月或者年的销售情况和平均值的关系。



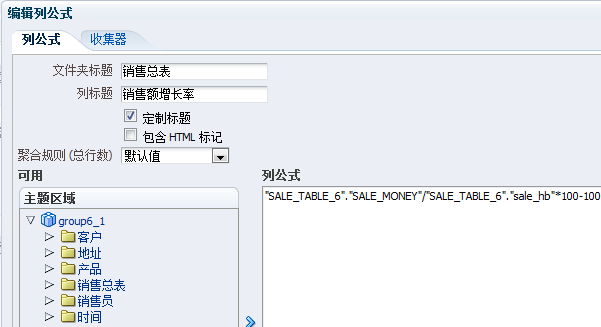


2）：同比环比列，度量值增长率列：

同比环比是实习中老师重点强调的两个基本的度量值，我们在导入表的过程中使用ago函数求得单个销售记录（销售额，销售量）的前一年或者前一个月的记录，把它记录下来作为一个新的度量值，可以很好地反映出来商品的销售趋势。

在这个部分中，我拓展了一个销售额增长率的属性列，即将环（同）比值与度量值作比并以百分数形式反映出来，这样可以看出我们的度量值较比之前的变化比例，我认为这是一个有用的度量值。

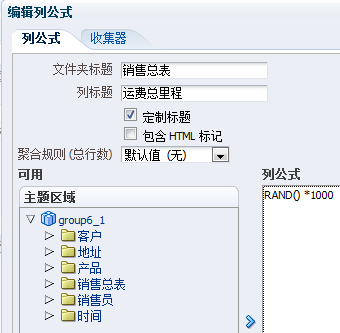




3）：运输总里程列：

由于我们组在数据导入的时候没有导入运输里程这个项目，所以我想设置一个运输里程项来反应运费的一些情况，让运费和运输方式，运输成本这三个量变得有意义，我在这里使用了rand函数在0~1000中随机生成一个数字将它赋给里程，只做了这个信息列。

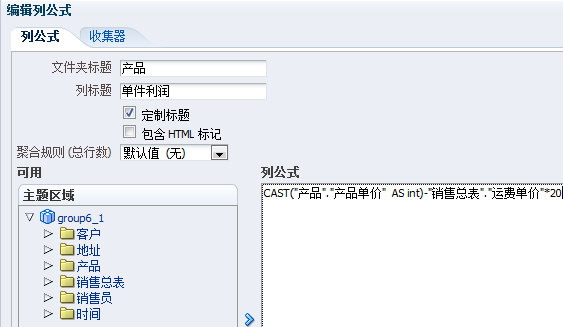




4）：各度量值的利润列：

产品总利润我们用产品单价减去产品成本之后乘以销售额，使用到了cast函数和—运算。

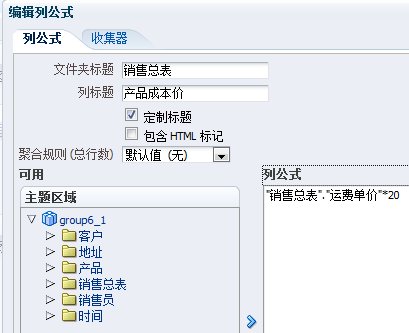


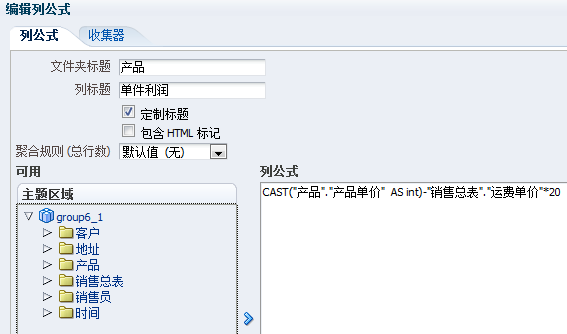


5）：产品成本价列：

同样由于之前的疏忽我们没有导入成本项，但是我们有一个多余的商品ID项可以修改，于是我们将它修改成为产品成本，于是我们就需要用cast函数对ID进行类型转换。



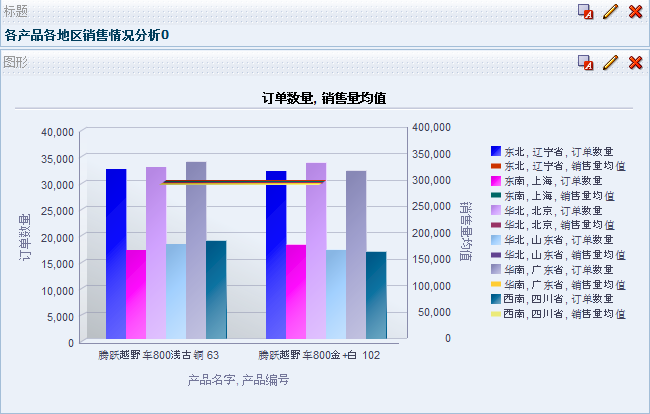


****

2.报表特色展示：

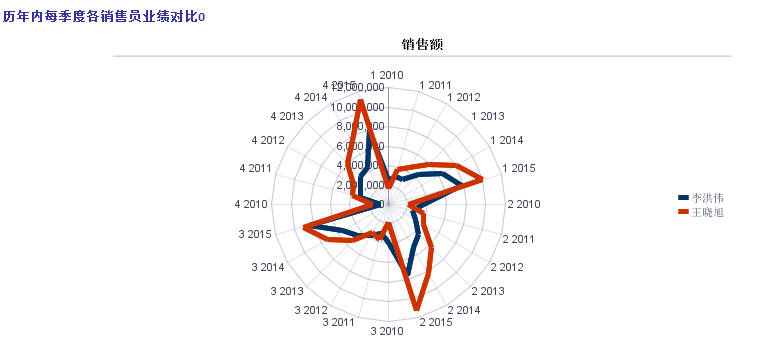
我在设计报表时主要考虑的是形式的一致性和对数据分析的合理性，所以我再设计统一的用折线图描述增长率的变化，有条形图描述数目的多少，用条线图描述度量值和均值之间的比较以及度量值之间的比较，用饼图描述各个值得占比情况，用雷达图描述数据的分布，用瀑布图描述利润和各个分利润之间的关系，用性能平铺图将度量值的总量展示出来。此后，由于我们的设计要求追求新意和美观程度，我有加入了计量表反应各个值的占比情况，加入了散点图描述数据的离散程度，加入格状图反应不同地区的销售情况对比，我最后增设了详细信息表，将各种没有办法从数据上体现出来的信息加入这个表当中，使数据库管理员能够了解除了数据之外的信息。

1）平均值比较条线图：



2）度量值雷达图：





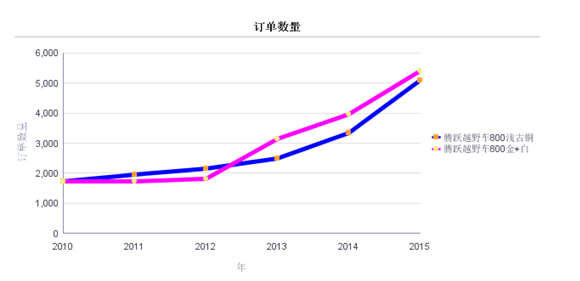
3）同（环）比增长率正负值条形图：

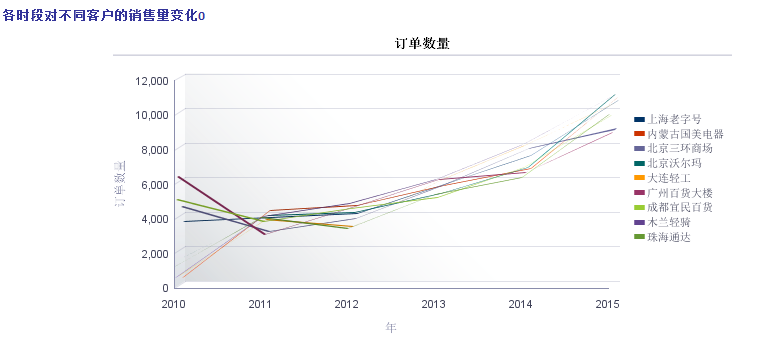


4）详细信息表：



5）对比折线图：

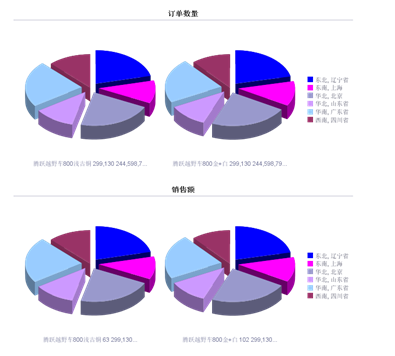




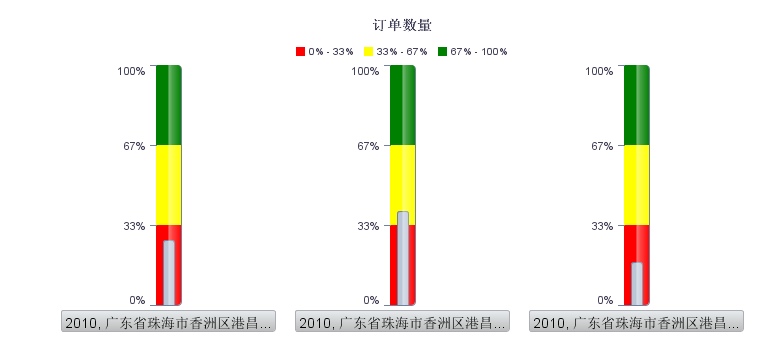
6）同（环）比增长条线图：



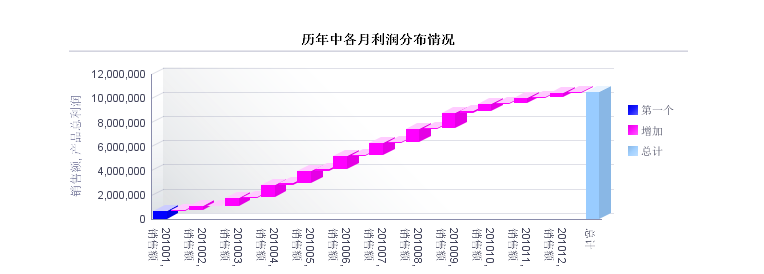
7）各部分占比，饼图：



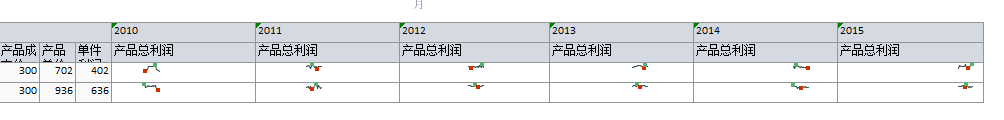
8）计量表：



9）反应月利润构成的瀑布图：



10）综合利润属性的格状图：



# 实习总结

## 实习内容的复杂性评价

本次实习的工程复杂性评价如表4.1所示：

表4.1 工程复杂性评价表

|  |  |  |
| --- | --- | --- |
|  | **复杂工程问题特征** | **问题描述及解决方案（方法）** |
| 1 | 必须运用深入的工程原理，经过分析才可能得到解决 | 首先我们学习了如何是对于数据仓库进行建模，我们才用星型模型，对原表中凌乱的数据进行系统的分析，从而最终确定下来我们需要尽力六个维度，因为这样可以最大化的体现我们的数据的价值。  此外，在对销售模型进行分析之前我们重点在学习SQL语句，这不但增强了我们的基本功，更让我们对于表中的数据有了深入的理解，让我们能够做到对自己统计的数据的意义有了明确的认识。 |
| 2 | 涉及多方面的技术、工程和其它因素，并可能相互有一定冲突 | 由于我们使用的OBIEE软件在导入目标表阶段只能单个小组进行操作，所以造成了资源利用上的冲突，当我们想要修改一些属性的时候没有办法说改就改，所以我们要动用自己的思维，在其他的步骤中进行修改，或者利用有限的资源事半功倍。 |
| 3 | 需要通过建立合适的抽象模型才能解决，在建模过程中需要体现出创造性 | 在建模中，周老师没有硬性要求我们建模过程中需要描述的数据和设计的维度，这需要我们在小组讨论的过程中自己去权衡，经过我们认真的商议，我们认为我们要建立一个六维度的表格，这样可以有效地反映出我们的数据的价值。  此外，在报表设计阶段我们的设计过程也是有很大自由度的，这就考研了我们的排版布局能力和审美意识，虽然我们组普遍以男生组成，审美意识匮乏，不过我们积极运用所学知识，在表格中拓展除了一些新的属性列，成为了我们的特色。 |
| 4 | 不是仅靠常用方法就可以完全解决的 | 在设计之初，我们难免有考虑不周全的时候，这时候就需要我们动用智慧而不是一味地推倒重来，我们在剪标的时候发现我们没有单价，利润这些属性，于是我们根据实际情况在创建报表的过程中添加了这些属性，避免了不必要的二次返工。  另外，再设计报表的时候，虽然我们无法像女同学的组那样做到十分的精致，不过我们另辟蹊径，设立了爆炸点，单品利润比这些有创意的小属性，使得我们的报表得到了增色。 |
| 5 | 问题中涉及的因素可能没有完全包含在专业工程实践的标准和规范中 | 在实际的生产生活中，对于一个企业的销售情况是一定要通过数据统计的形式给出的，这样才能直观地反应出来数据的离散度和对比度，在实际生产中，不会每个企业员工都要去研究销售的数据，所以数据分析师就是将数据统计起来，制作成更加直观地报表，给企业中的其他人员带来方便，所以我们的生产实习的方向是一个非常现实而又具体的方向。是一定会包含在专业工程时间的标准和规范中的。 |
| 6 | 问题相关各方利益不完全一致 | 在数据采集阶段，如果注重数据的全面性那么必然会损失简洁性，反之，如果一味看重简洁性，那么必然导致关键数据的缺失，在这一点上我们一直在讨论和修改中不断完善，这种急急急的讨论过程也是贯穿整个生产实习始终的。 |
| 7 | 具有较高的综合性，包含多个相互关联的子问题 | 制作报表是我们的最终目标，然然在整个制作过程中我们总共接触了四个软件，一种语言，三种建模方式，可以说综合性还是很强的。  在生产实习的过程中，老师一直在丰富我们对于不同软件的应用能力，我们也在这个过程中收益匪浅。 |

## 实习体会、收获与建议

在这短短三十天的时间里，通过我们日复一日的工作让我对BI，这个我曾不熟悉的行业有了一个初步的认识，当我的报表制作完成的时候，我了解了一个基本的BI工程的流程，虽然我们的项目是排除了许多的客观的复杂因素的，不过我想今后在生产生活中如果面对困难我们还是要一样的解决。

实习的内容固然重要，但是更加重要的我想应该是培养这种自学的能力，在实习的时候周老师反复告诉我们，不会的地方先百度，实在不会了在找老师解决。在发现问题和解决问题的过程中，我们收获的不仅仅是对SQL语句，数据仓库建模这些知识点的理解，而是一种宏观的解决问题的办法。这种办法就是自学。在以后的生产生活中，能带给自己指导的人将越来越少，此外，还有很多的新知识等着我们去挖掘，自学能力就是一个人的综合能力的良好体现，所以如果说实习意在教授我们知识，不如说实习其实是在锻炼我们自学。

此外，在实习中的团队协作能力也尤为重要，虽然这次的任务是一个小任务，但是可想而知当以后经历重大的任务的时候我们需要组成团队才能完成任务，在实习的过程中，周老师反复给我们强调小组的重要性，让我们首先明确我们是一个团体，而不是单打独斗的个人这一点，我想这也是今后在企业中将让我们受益的一个重要的素质吧。

从接到任务到完成任务，我们经历了三十天的时间，然而在这三十天里我们并不是一直在制作报表，我们刚开始的时候丰富了我们原有的SQL知识，接着我们学习了数据仓库应该如何建模，然后我们尝试把数据导入新表格，最后才是制作报表，其实如果是按照工程效率的角度，我们完全可以直接在OBIEE中完成这些修改，然而我们却前后经历了四种软件，三大块知识点，老师如此设计的目的我想就是希望我们能够多了解一些BI相关的知识，多掌握几种软件的使用方法，无论今后从事什么行业，技不压身的道理是一定会体现出来的，我们要做的就是在学生阶段不断地丰富自己，像一颗海绵一样拼命地吸取知识，因为我们不知道在未来的工作中我们会经历什么样的挑战，我们会用到什么样的知识，在一次次反复学习的过程中，我们终将丰富自己的阅历，提升自己的工作本领。

在大学里经历了三年的学习生活后，面对这次模拟企业环境的生产实习，我感受到了学校和单位的不一样，在学校里你总是要学习那么多的知识不过一定是要到期末才会有一张卷纸来评价你学的好坏，不过在生产实习中你学的越快，你完成的越多，你就比别人做的越出色，等到我们真正走上岗位的时候，你就比别人赚的越多，你就别别人活得越好。这虽然世俗，但就是现实。在今后的工作中容不得我们半点的懈怠，没有平时不努力期末突击也能顺利了事的道理，只有一份付出一份回报的等价交换。

最后我想感叹计算机这门科学，在如此短时间里竟然已经发展出了如此复杂的分支，虽然同样是SQL数据库的知识，但是还是和我们在学校学习的知识完全不同，我们身处在这个日新月异的行业中，它引领着时代的进步，所以我们更需要不断完善自己，方能跟上这日新月异的步伐。也许处理计算机行业，其他的行业也是如此，所以我想，我们不应该去考虑自己要学什么，不要学什么，在今天这么多变的环境中也许你今天学完的知识明天就被更新更好的知识所取代了，也许你已经熟练掌握的技能还没来得及应用就已经被时代抛弃了，但是，这并不能影响你继续探索的脚步，因为学习是让你时刻保持进取的最好的强心剂！

生产实习的结束也意味这大学的尾声，短短三十天让我遇见了非常尽职尽责的老师和友善的同学们，在未来的日子里，希望老师同学们都生活顺利，梦想成真，也希望自己能够继续保持着一个求实进取的心，在今后的生活中更进一步！

# 参考文献

1. 米尔顿（MichealMilton），深入浅出数据分析. 美国：电子工业出版社，2012.
2. 西格兰，哈梅巴赫，数据之美.南京:东南大学出版社,2010.
3. 申德荣.数据库原理课件.沈阳:.东北大学出版社,2017.

**附录 实习日记**

|  |  |
| --- | --- |
| 7月3日 | 上午听课，周老师介绍了生产实习的具体内容，我们需要先掌握SQL的知识，所以我们之前的课程重点在于掌握SQL，熟悉软件。 |
| 7月4日 | 上午听课，练习SQL知识，熟悉PL/SQL软件的引用。 |
| 7月5日 | 第一次讲座。 |
| 7月6日 | 上午听课，练习SQL知识，熟悉PL/SQL软件的引用。 |
| 7月7日 | 上午听课，练习SQL知识，熟悉PL/SQL软件的引用。 |
| 7月8日 | 休息 |
| 7月9日 | 休息 |
| 7月10日 | 上午听课，练习SQL知识，熟悉PL/SQL软件的引用。 |
| 7月11日 | 第二次讲座 |
| 7月12日 | 上午上课，介绍数据仓库的有关知识，同时接触了新的软件。 |
| 7月13日 | 上午上课，分析数据，继续熟练SQL |
| 7月14日 | 上午上课，分析数据，进行建模 |
| 7月15日 | 休息 |
| 7月16日 | 第三次讲座，下午上课，进行建模 |
| 7月17日 | 全天上课，实现自己的建模过程，建立目标表 |
| 7月18日 | 参观东软，下午上课，建立目标表 |
| 7月19日 | 全天上课，实现自己的建模过程，建立目标表 |
| 7月20日 | 全体上课，学习了如何像目标表中添加数据 |
| 7月21日 | 休息 |
| 7月22日 | 全天上课，像目标表中添加数据 |
| 7月23日 | 全天上课，像目标表中添加数据，完善表格的属性 |
| 7月24日 | 全天上课，学习了如何将目标表与OBIEE相连接 |
| 7月25日 | 全天上课，同时将我们的目标表连接到OBIEE上 |
| 7月26日 | 全天上课，制作报表 |
| 7月27日 | 全天上课，制作报表，完善报表 |
| 7月28日 | 全天上课，制作报表，完善报表 |
| 7月29日 | 全天上课，制作报表，完善报表，小组内进行整理汇总 |
| 7月30日 | 上午对项目进行了全面检查和完善，并提交给老师进行程序验收。 |
| 9月1日 | 提交实习报告 |