第二章 营销零售

2017年5月25日

9:31

**一.四步维度设计过程**

**1.选取业务处理过程:**

业务处理过程是源数据系统的自然业务活动;

分析活动的业务度量值来自业务处理过程,选取适当的业务处理过程,最高的效率是听取用户的意见;

建立的维度模型是业务处理过程,比如订单数据;

而不是为要存取订单的销售和市场部门,建立单独的维度模型;

这样就可以在公司范围内,经济的提供一致的数据,对于不同的部门;

如果建立的维度模型和部门捆绑在一起的,则维度可能有不同的标记;

数据有拷贝.多重的数据流向维度模型,不一致性就会脆弱;

确保一致性的最佳方法是,对数据进行一次性发布,而且也减少了ETL的开发量;

**2.定义业务的处理粒度**

粒度定义:给出事实表的行的实际含义给出明确的说明;

就是事实表度量值的细节所达到的程度;

团队中的表的粒度定义一定要一致;

**3.选定用于事实表行(原子粒度)的维度**

怎么描述业务处理过程中的度量值;

维度其实就是, 每个度量取单一值而代表的所有可能情况的丰富描述(列越多越好);

维度的选定,就是使维度表丰满起来的离散的文本属性;

厂家的维度:日期,产品,顾客,事务类型和状况;

**4.确定用于形成每个 事实表行的数字型事实**

要对什么内容进行评测

以用户对业务的理解作为确定维度模型所需维度,事实内容的依据

**二.零售实例研究**

**1.业务处理**

通过对业务需求的理解 对可用数据的理解

组合起来而确定建模业务的处理内容;

用户希望更好的理解POS系统记录的顾客购买情况;

选取的业务处理过程就是POS零售业务;

**2.定义粒度**

在维度模型中给出何种详细程度细节内容;

优先考虑,最具原子性的信息而开发维度模型;

原子粒度能确切了解更多的事情, 而且确切知道的事情,就能转换为维度

最佳粒度是POS事务的单个分列项,不是为了查询低层面的行,而是为了以精确的方式对细节的知识进行查询

**3.选取维度**

粒度选定,确切知道的信息(时期,产品,商店)的维度也就是随之而且度确定下来了

**4.确定事实**

粒度定义是支撑点

事实对于粒度必须是真实的: 销售量,单价,销售额.产品的标准出厂价

而且可以直接被报表一致性的引用

诸如百分比这种非加性,分子,分母都存放在事实表中;

可以通过BI工具求出比率

单价也是非加型事实

**5.维度属性**

***日期维度***

10年的每一天进行存储;

每列有行所代表的特定日期定义,比如星期,星期,月,季度,节假日

周末指示符:容易实现平日假期和周末假期进行比较这样的应用;

销售时令:可以是圣诞节,感恩节,复活节,情人节,独立日,或者不是;

重大事件:超级活动日,其他影响事件等特殊外部事件;

***产品维度:***

***商场维度***

商场维度是基本的地理维度:

***促销维度***

临时降价,报纸广告和优惠卷;

***退化维度***

***购物常客维度表:***

**二.市场篮子分析**

一起售出的产品的组合分析.

数据挖掘:相似成组.

市场篮子事实表: