



Data Warehouse (数据仓库)

数据仓库是一个各种数据（包括历史数据和当前数据）的中心存储系统，是BI（business intelligence，商业智能）的核心部件。

这里所谈的数据包括来自企业业务系统的订单、库存、交易账目、客户和供应商等来自企业所处行业和竞争对手的数据以及来自企业所处的其他外部环境中的各种数据。

Business Intelligence (商业智能)

商业智能通常被理解为将企业中现有的数据转化为知识，帮助企业做出明智的业务经营决策的工具。

为了将数据转化为知识，需要利用数据仓库、联机分析处理（OLAP）工具和数据挖掘等技术。



OLAP (online analytical processing)

OLAP (online analytical processing) 是一种软件技术，它使分析人员能够迅速、一致、交互地**从各个方面观察信息**，以达到深入理解数据的目的。从各方面观察信息，也就是从不同的维度分析数据，因此OLAP也成为**多维分析**。

订单id	下单地区	下单品类	下单时间	订单金额
1001	华北	电子	12月	456
1002	华东	食品	11月	489
1003	西南	居家	2月	491
1004	东北	电子	4月	659
1005	西北	宠物	11月	369
1006	华北	食品	12月	159

分析订单金额**总和**的规律，需要**从不同的角度**去观察

角度1：地区

角度2：品类

角度3：时间

角度4：地区， 品类

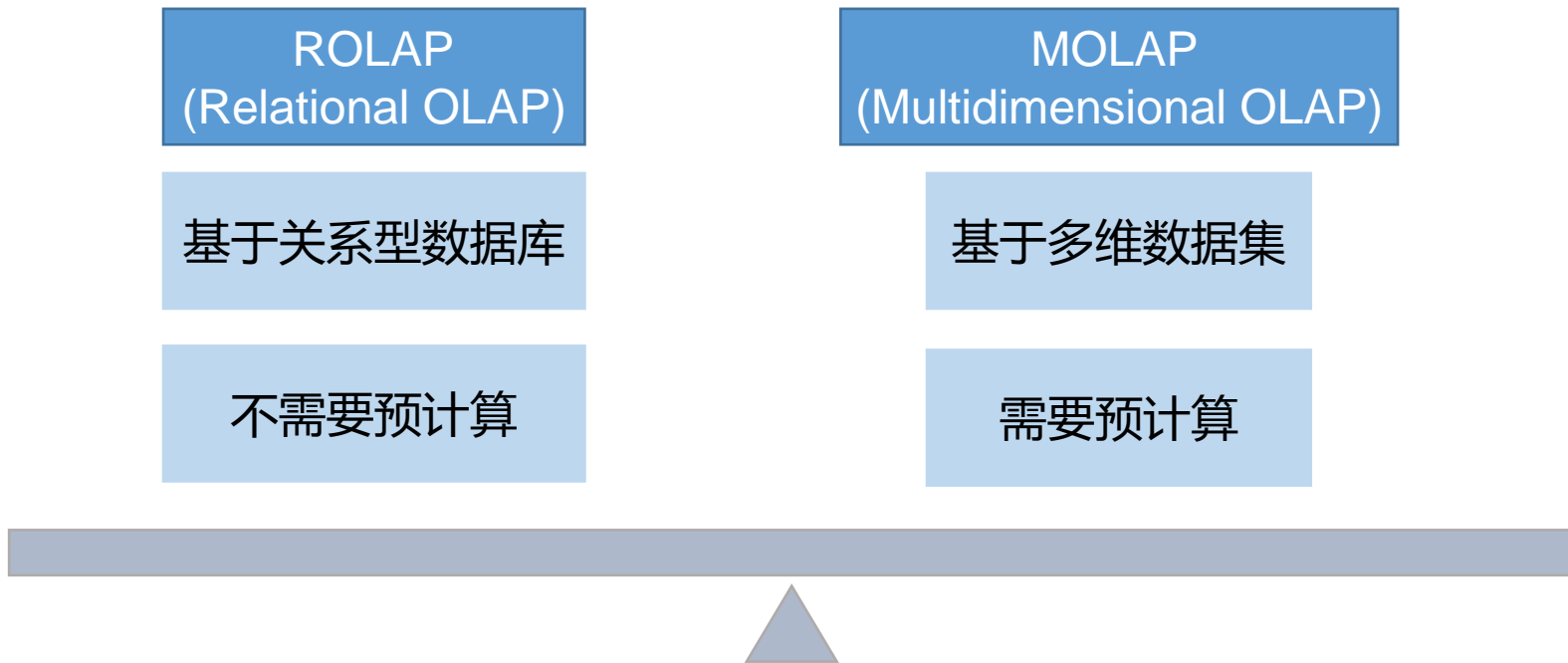
角度5：地区， 时间

角度6：品类， 时间

角度7：地区， 品类， 时间



OLAP 类型

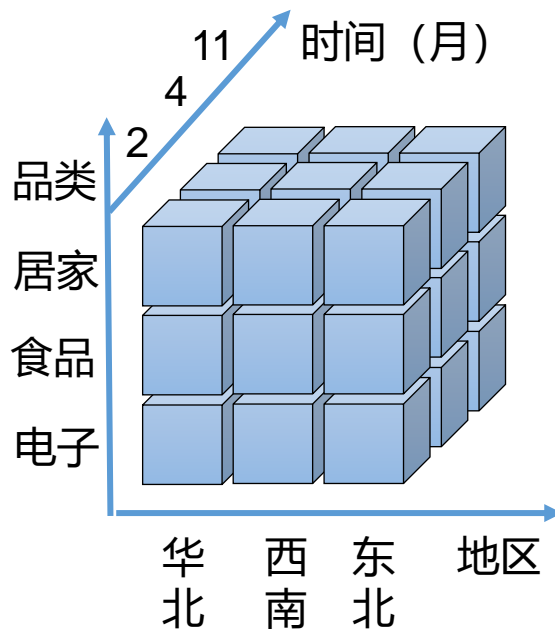




OLAP Cube

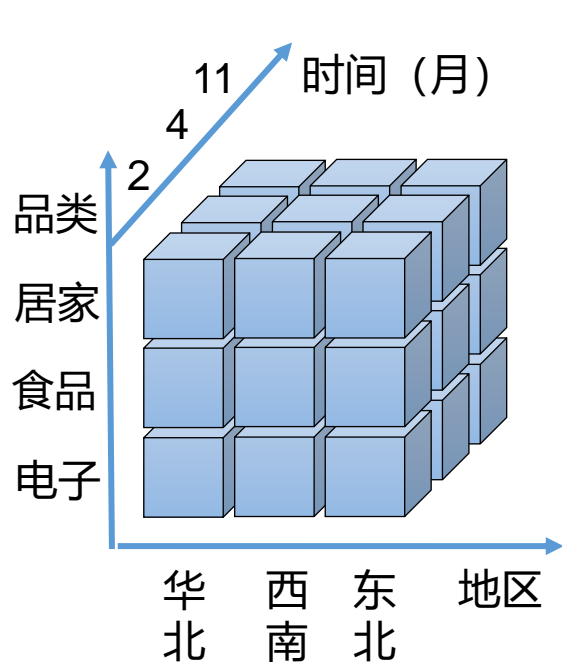
MOLAP基于多维数据集，一个多维数据集称为一个OLAP Cube

订单id	下单地区	下单品类	下单时间	订单金额
1001	华北	电子	12月	456
1002	华东	食品	11月	489
1003	西南	居家	2月	491
1004	东北	电子	4月	659
1005	西北	宠物	11月	369
1006	华北	食品	12月	159

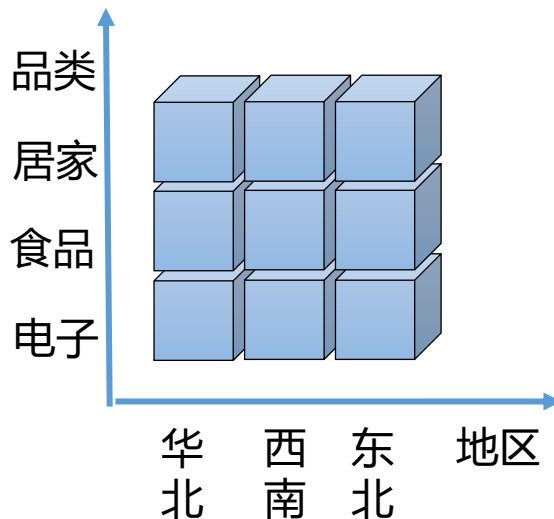




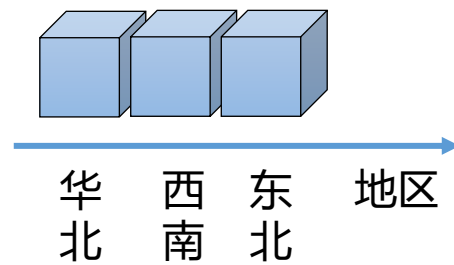
Cube & Cuboid



角度7: 地区, 品类, 时间



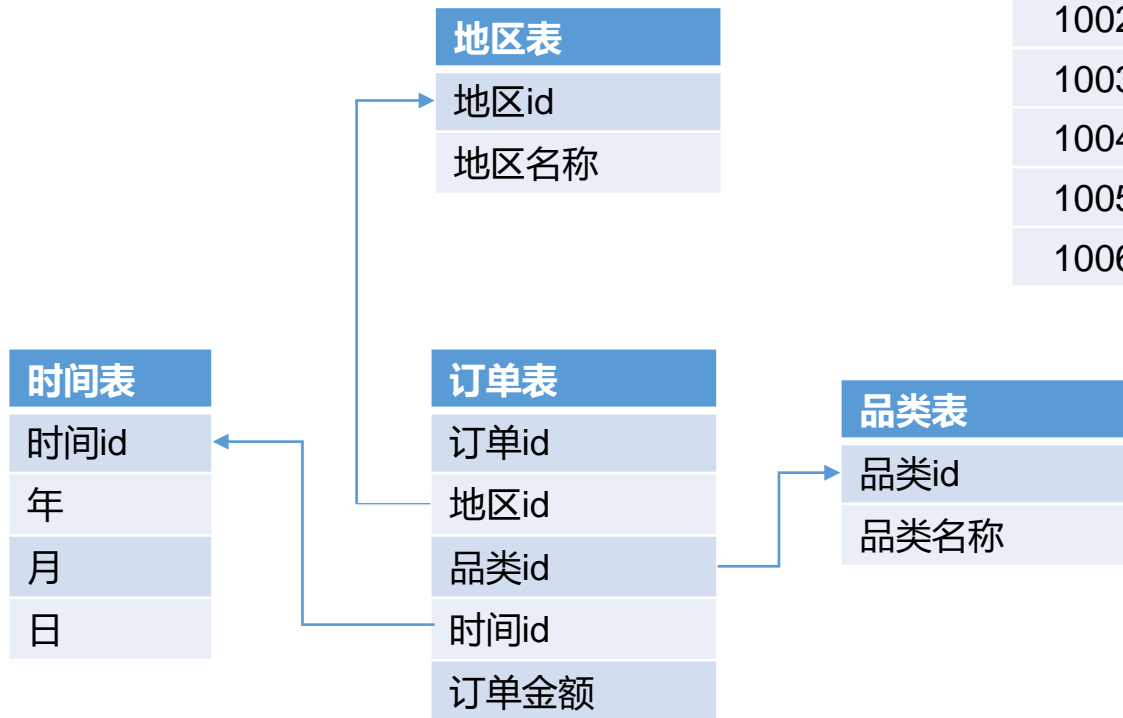
角度4: 地区, 品类
角度5: 地区, 时间
角度6: 品类, 时间



角度1: 地区
角度2: 品类
角度3: 时间



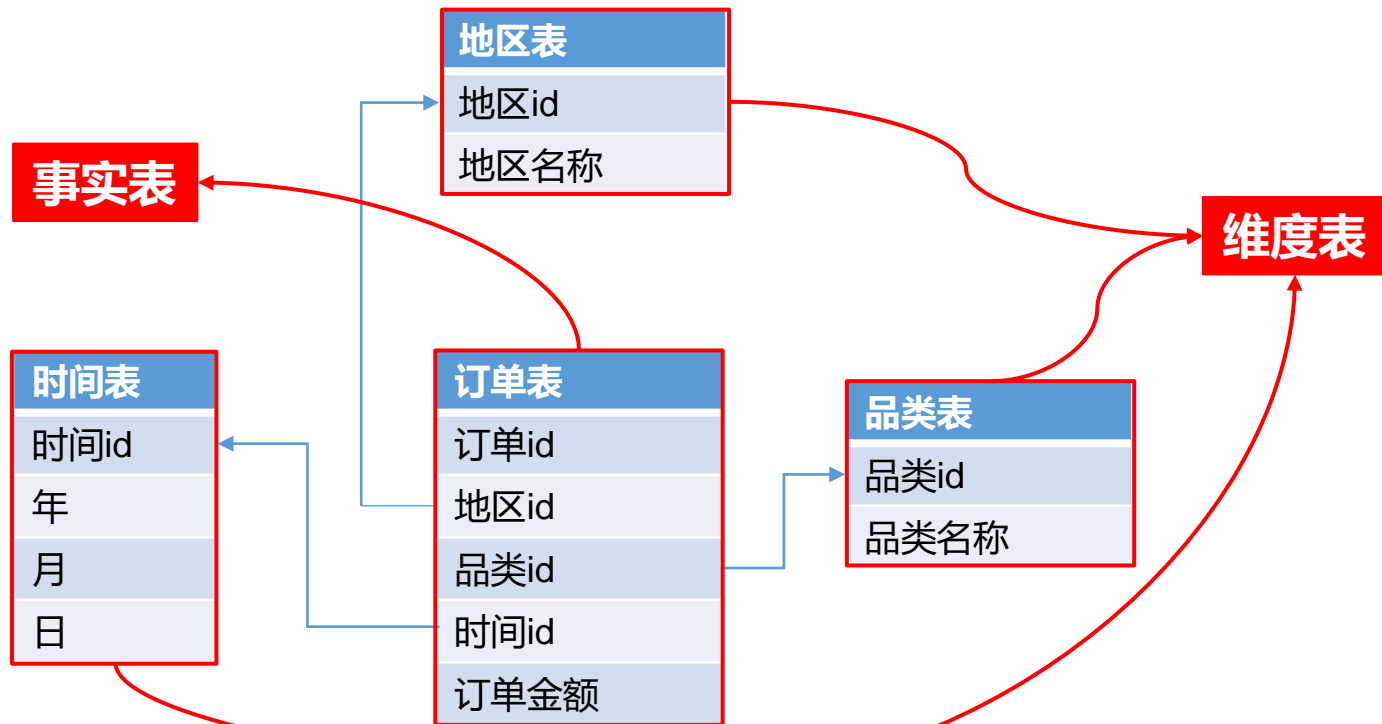
Star Schema (星型模型)



订单id	下单地区	下单品类	下单时间	订单金额
1001	华北	电子	12月	456
1002	华东	食品	11月	489
1003	西南	居家	2月	491
1004	东北	电子	4月	659
1005	西北	宠物	11月	369
1006	华北	食品	12月	159



Fact Table (事实表) & Dimension Table (维度表)





Dimension (维度) & Measure (度量)

订单id	维度			度量
	下单地区	下单品类	下单时间	订单金额
1001	华北	电子	12月	456
1002	华东	食品	11月	489
1003	西南	居家	2月	491
1004	东北	电子	4月	659
1005	西北	宠物	11月	369
1006	华北	食品	12月	159

维度：分析数据的数据的角度

度量：被分析的指标