



# Managing Data & Databases

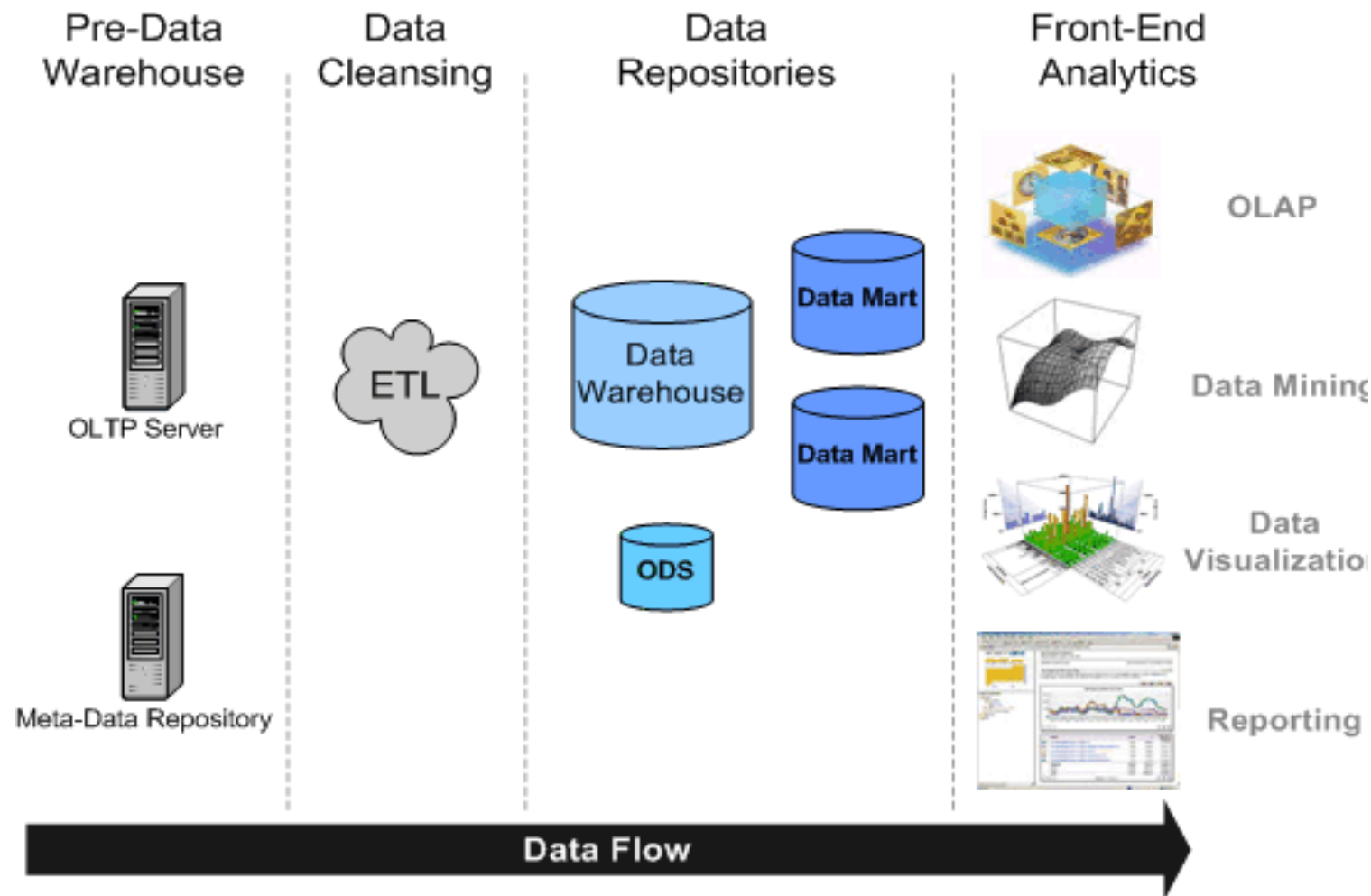
Session 21

Dimensional Data, Data Warehouses and BI

## Some BI jargon

- ETL
- OLAP
  - Dicing, Slicing, Drilling, Pivoting
- Data Warehouse
- Data Mart
- Cube
- Star Schema
- Snowflake Schema

## What happens after ETL?



# Data warehouse and Data mart

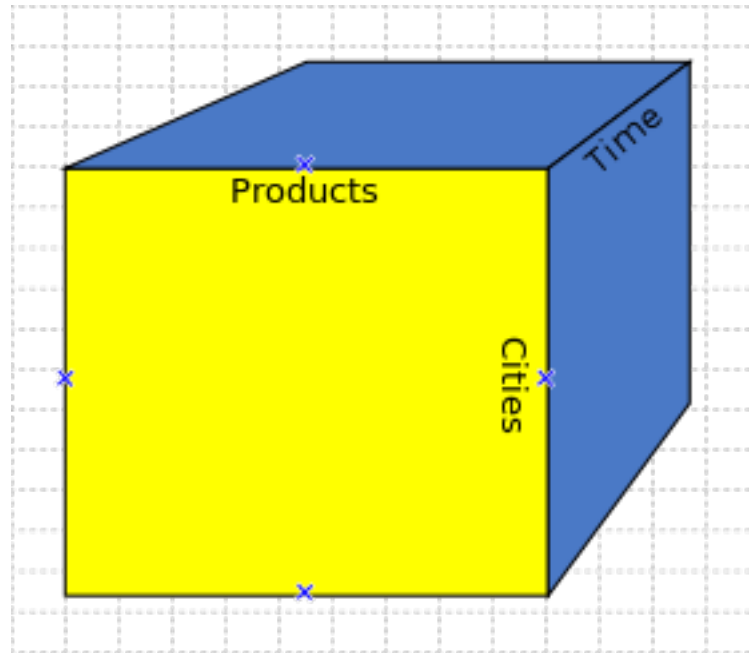
- Data warehouse

- Integrates all available and analyzable data of the firm
- On different subjects
- With complete detail
- For the purpose of online analytical processing, data mining, reporting or visualization

- Data marts

- Subject-specific, analysis-ready pockets of data designed for multidimensional queries and drill-downs

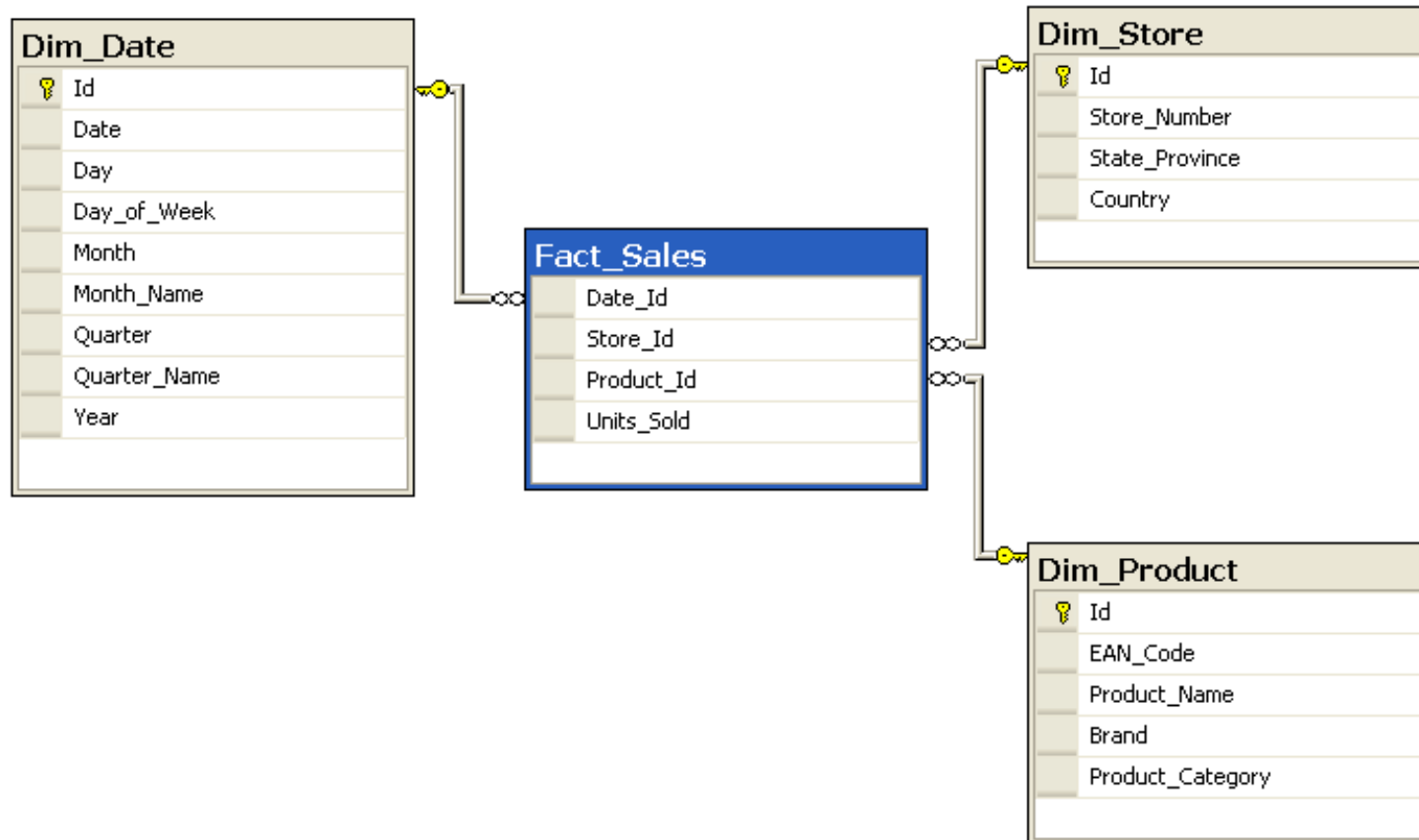
# Cubes



# A multidimensional query

How many TV sets have been sold, for each brand and country, in 1997?

# Star Schema



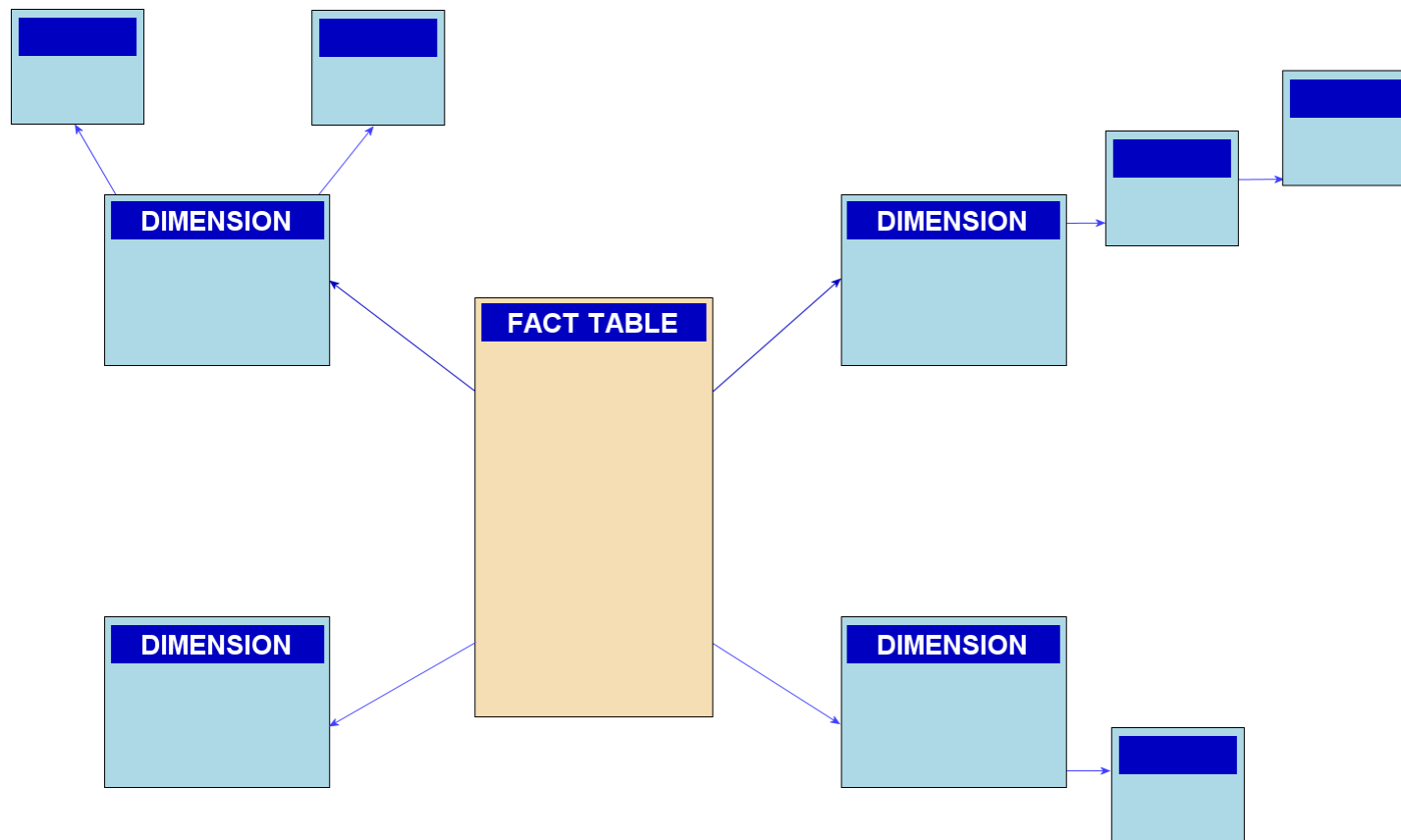
# Star Schema

```
SELECT
    P.Brand,
    S.Country AS Countries,
    SUM(F.Units_Sold)
FROM Fact_Sales F
INNER JOIN Dim_Date D    ON F.Date_Id = D.Id AND D.Year = 1997
INNER JOIN Dim_Store S   ON F.Store_Id = S.Id
INNER JOIN Dim_Product P ON F.Product_Id = P.Id AND P.Product_Category = 'tv'

GROUP BY
    P.Brand,
    S.Country
```



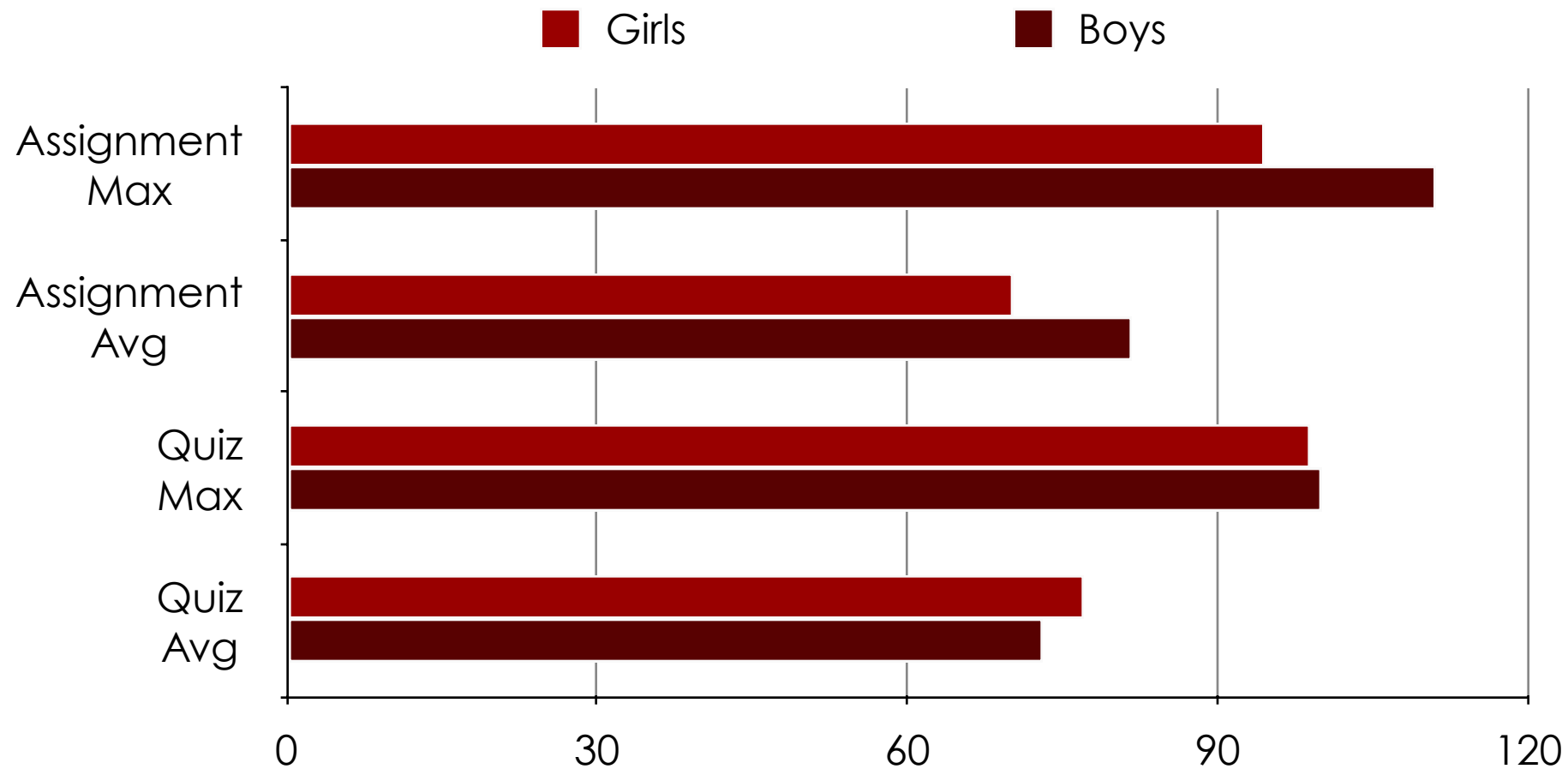
# Snowflake Schema



# Pivot Tables

A3		fx	Sum of Amount							
	A	B	C	D	E	F	G	H	I	J
1	Category	(All)								
2										
3	Sum of Amount	Product								
4	Country	Apple	Banana	Beans	Broccoli	Carrots	Mango	Orange	Grand Total	
5	Australia	20634	52721	14433	17953	8106	9186	8680	131713	
6	Canada	24867	33775		12407		3767	19929	94745	
7	France	80193	36094	680	5341	9104	7388	2256	141056	
8	Germany	9082	39686	29905	37197	21636	8775	8887	155168	
9	New Zealand	10332	40050		4390			12010	66782	
10	United Kingdom	17534	42908	5100	38436	41815	5600	21744	173137	
11	United States	28615	95061	7163	26715	56284	22363	30932	267133	
12	Grand Total	191257	340295	57281	142439	136945	57079	104438	1029734	
13										
14										

## A grade overview...



Suggestions?

Evaluation

Quiz Question Submission