Data Warehouse Schema

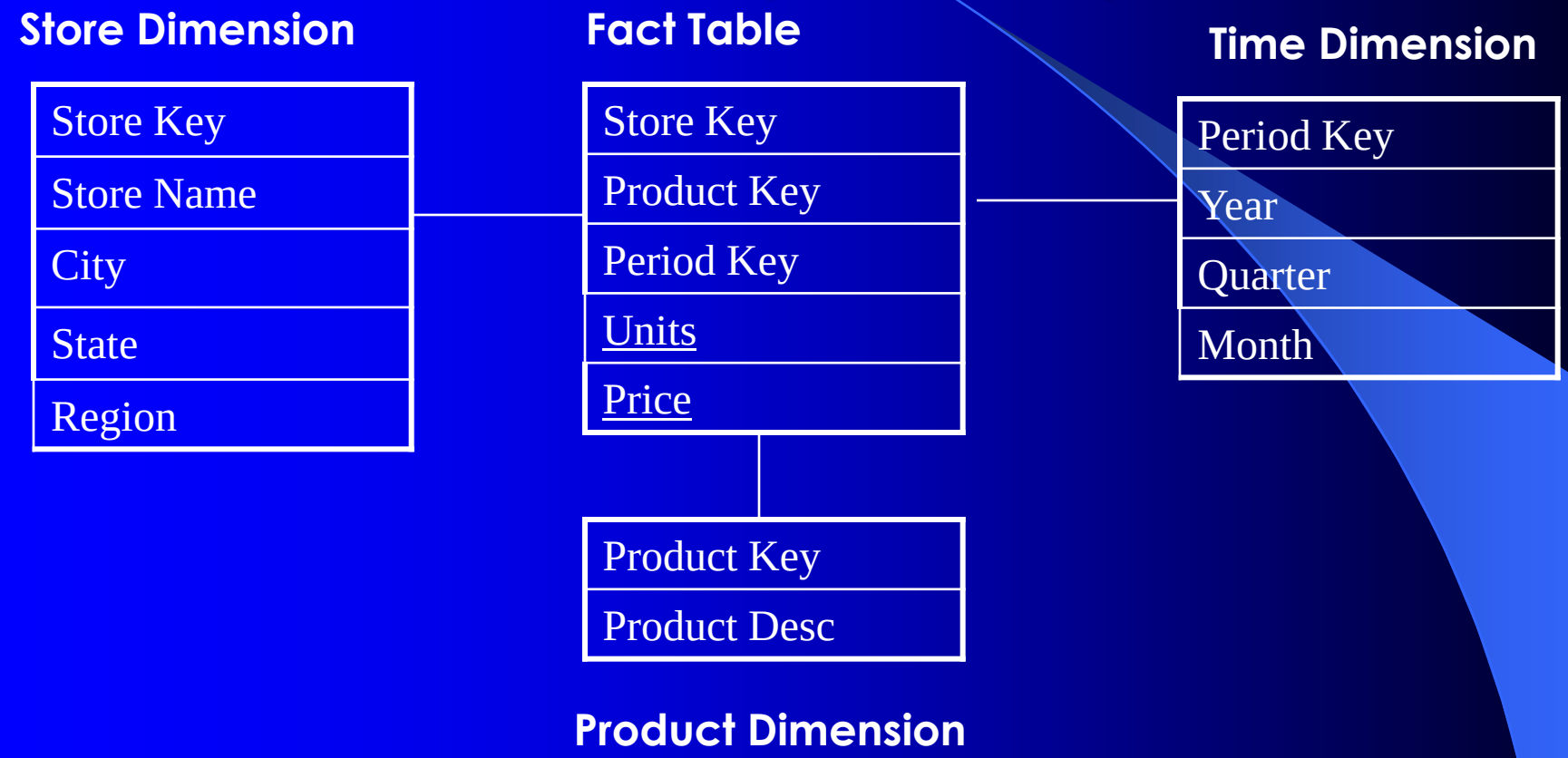
Data Warehouse Schema

- Star Schema
- Snowflake Schema
- Fact Constellation Schema

Star Schema

- A single,large and central fact table and one table for each dimension.
- Every fact points to one tuple in each of the dimensions and has additional attributes.
- Does not capture hierarchies directly.

Star Schema (contd..)

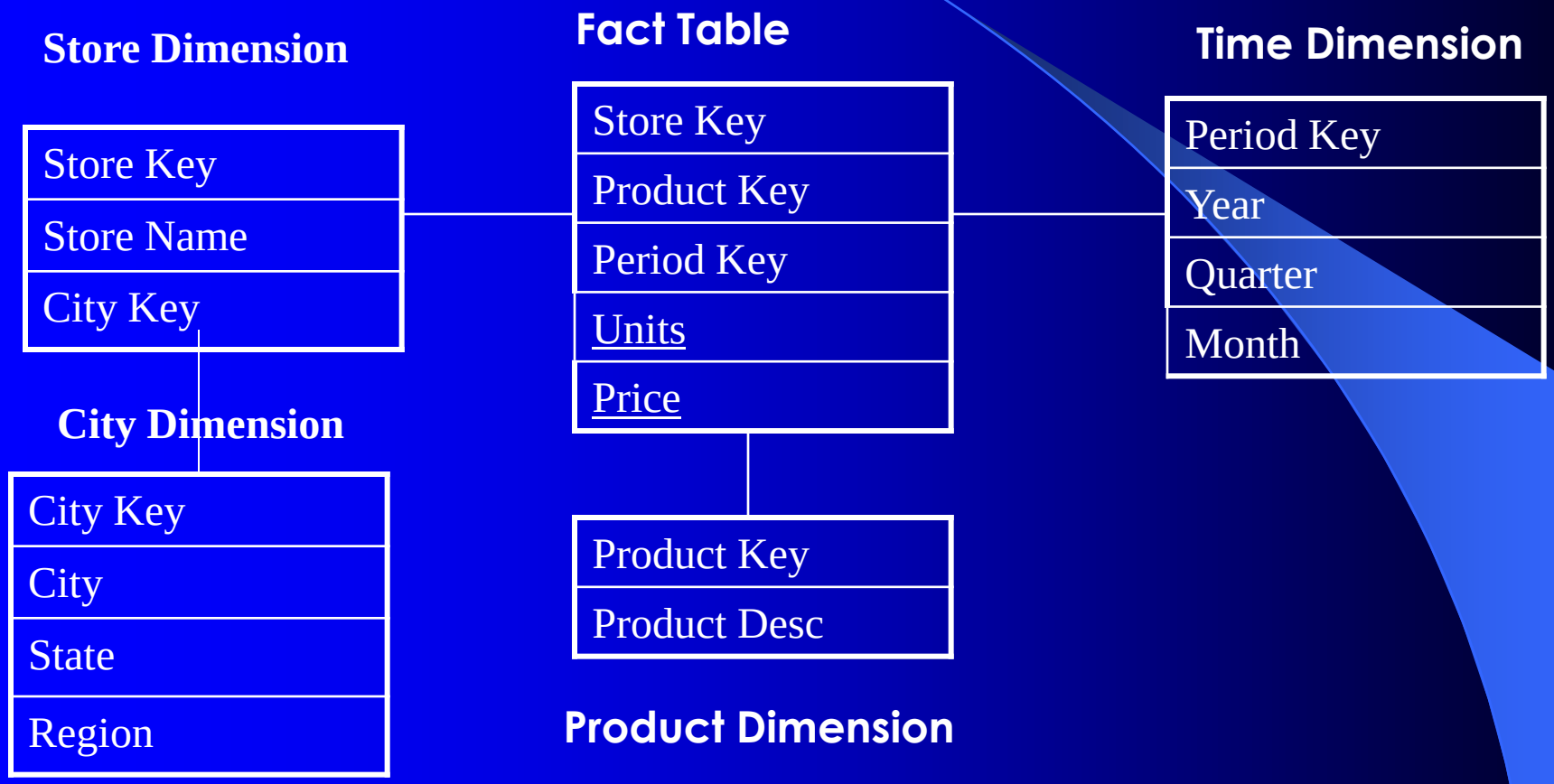


Easy to understand, easy to define hierarchies, reduces no. of physical joins.

SnowFlake Schema

- Variant of star schema model.
- A single,large and central fact table and one or more tables for each dimension.
- Dimension tables are normalized i.e. split dimension table data into additional tables

SnowFlake Schema (contd..)

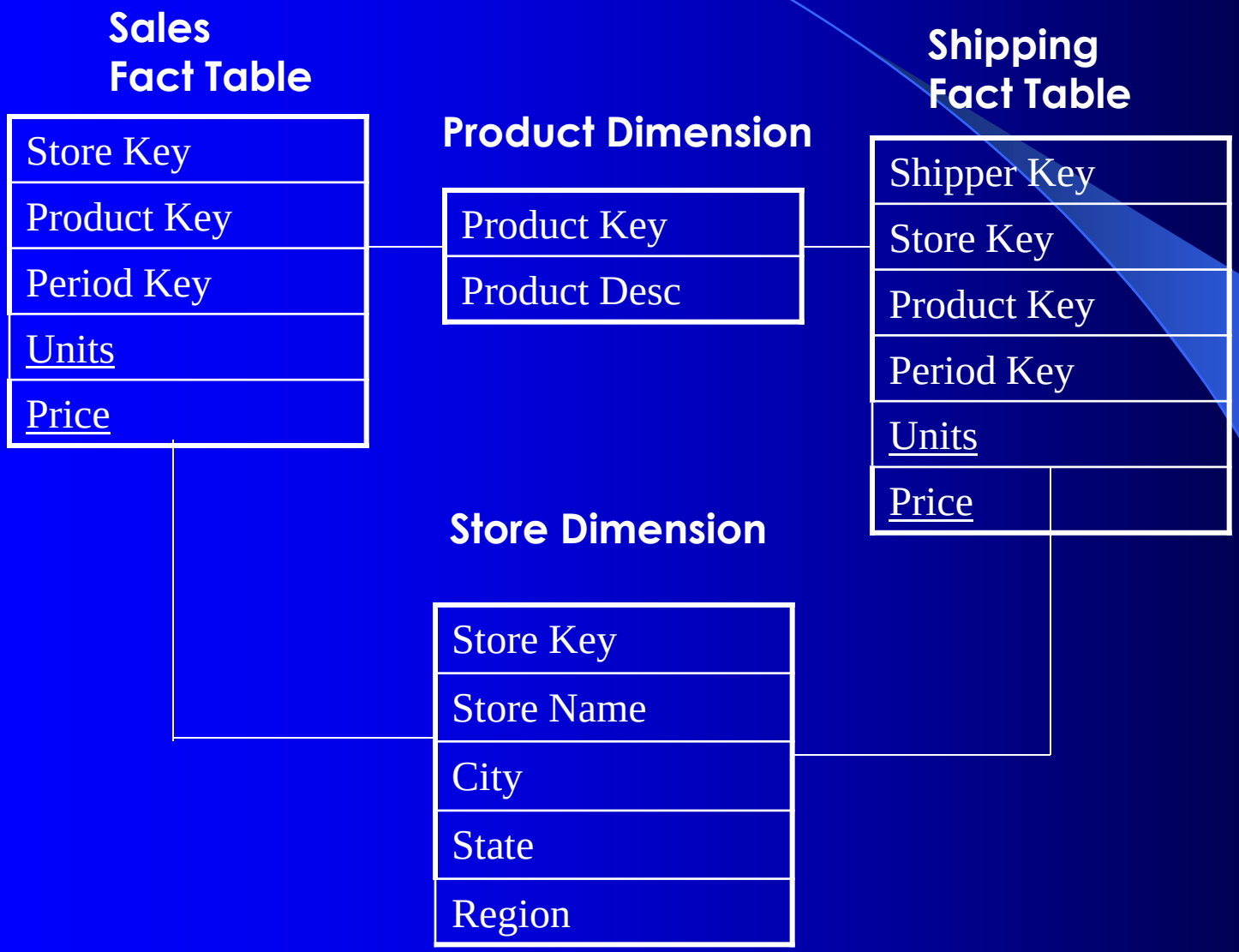


Drawbacks: Time consuming joins,report generation slow

Fact Constellation

- Multiple fact tables share dimension tables.
- This schema is viewed as collection of stars hence called galaxy schema or fact constellation.
- Sophisticated application requires such schema.

Fact Constellation (contd..)



Building Data Warehouse

- Data Selection
- Data Preprocessing
  - Fill missing values
  - Remove inconsistency
- Data Transformation & Integration
- Data Loading

Data in warehouse is stored in form of fact tables and dimension tables.

Case Study

- Afco Foods & Beverages is a new company which produces dairy,bread and meat products with production unit located at Baroda.
- There products are sold in North,North West and Western region of India.
- They have sales units at Mumbai, Pune , Ahemdabad ,Delhi and Baroda.
- The President of the company wants sales information.

Sales Information

Report: The number of units sold.

113

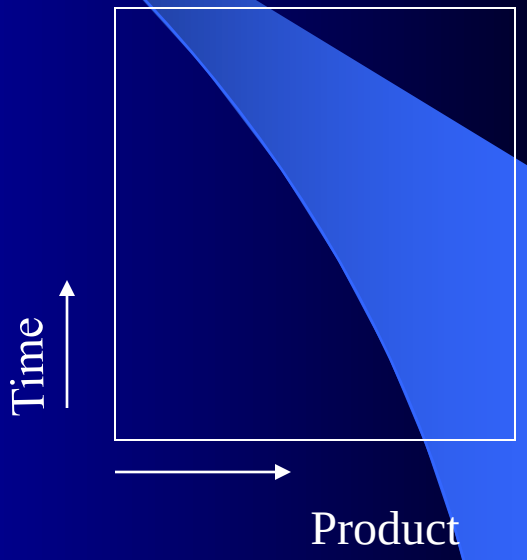
Report: The number of units sold over time

January	February	March	April
14	41	33	25

Sales Information

Report : The number of items sold for each product with time

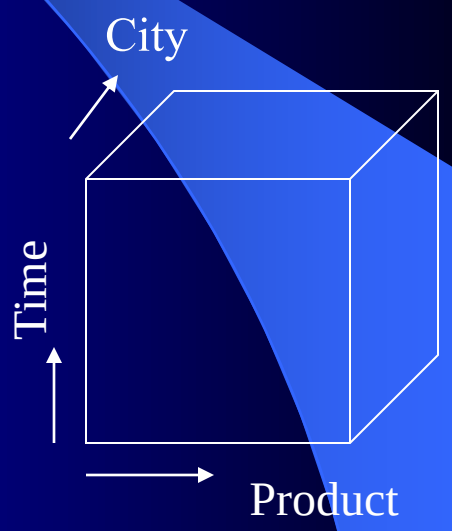
	Jan	Feb	Mar	Apr
Wheat Bread			6	17
Cheese	6	16	6	8
Swiss Rolls	8	25	21	



Sales Information

Report: The number of items sold in each City for each product with time

		Jan	Feb	Mar	Apr
Mumbai	Wheat Bread			3	10
	Cheese	3	16	6	
	Swiss Rolls	4	16	6	
Pune	Wheat Bread			3	7
	Cheese	3			8
	Swiss Rolls	4	9	15	



Sales Information

Report: The number of items sold and income in each region for each product with time.

		Jan		Feb		Mar		Apr	
		Rs	U	Rs	U	Rs	U	Rs	U
Mumbai	Wheat Bread					7.44	3	24.80	10
	Cheese	7.95	3	42.40	16	15.90	6		
	Swiss Rolls	7.32	4	29.98	16	10.98	6		
Pune	Wheat Bread					7.44	3	17.36	7
	Cheese	7.95	3					21.20	8
	Swiss Rolls	7.32	4	16.47	9	27.45	15		

Sales Measures & Dimensions

- Measure – Units sold, Amount.
- Dimensions – Product,Time,Region.

Sales Data Warehouse Model

Fact Table

City	Product	Month	Units	Rupees
Mumbai	Wheat Bread	January	3	7.95
Mumbai	Cheese	January	4	7.32
Pune	Wheat Bread	January	3	7.95
Pune	Cheese	January	4	7.32
Mumbai	Swiss Rolls	February	16	42.40

Sales Data Warehouse Model

City_ID	Prod_ID	Month	Units	Rupees
1	589	1/1/1998	3	7.95
1	1218	1/1/1998	4	7.32
2	589	1/1/1998	3	7.95
2	1218	1/1/1998	4	7.32
1	589	2/1/1998	16	42.40

Sales Data Warehouse Model

Product Dimension Tables

Prod_ID	Product_Name	Product_Category_ID
589	Wheat Bread	1
590	White Bread	1
288	Coconut Cookies	2

Product_Category_Id	Product_Category
1	Bread
2	Cookies

Sales Data Warehouse Model

Region Dimension Table

City_ID	City	Region	Country
1	Mumbai	West	India
2	Pune	NorthWest	India

Sales Data Warehouse Model