

Projeto Relacional do Data Warehouse

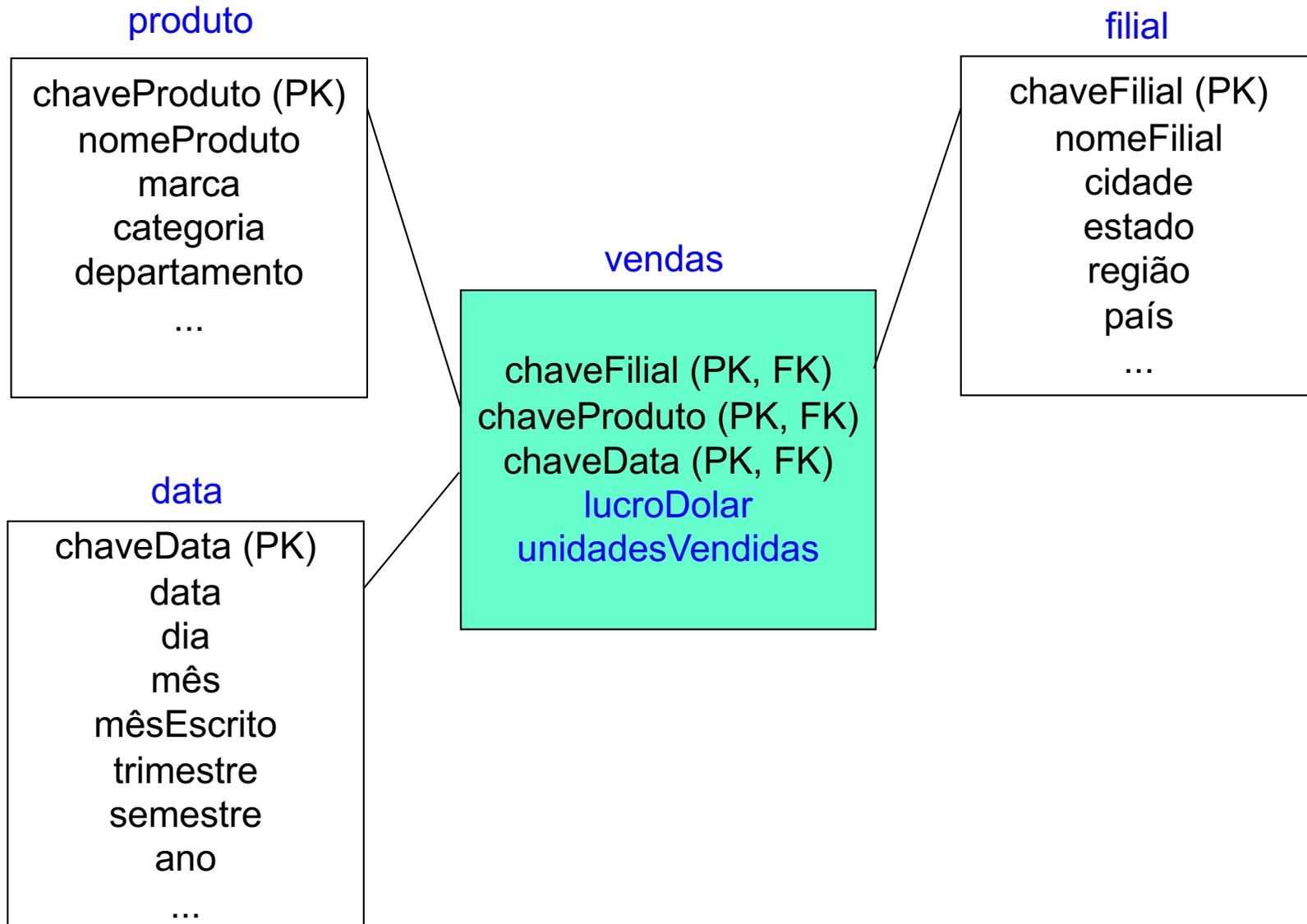
Processamento Analítico de Dados
Profa. Dra. Cristina Dutra de Aguiar

Sistemas ROLAP

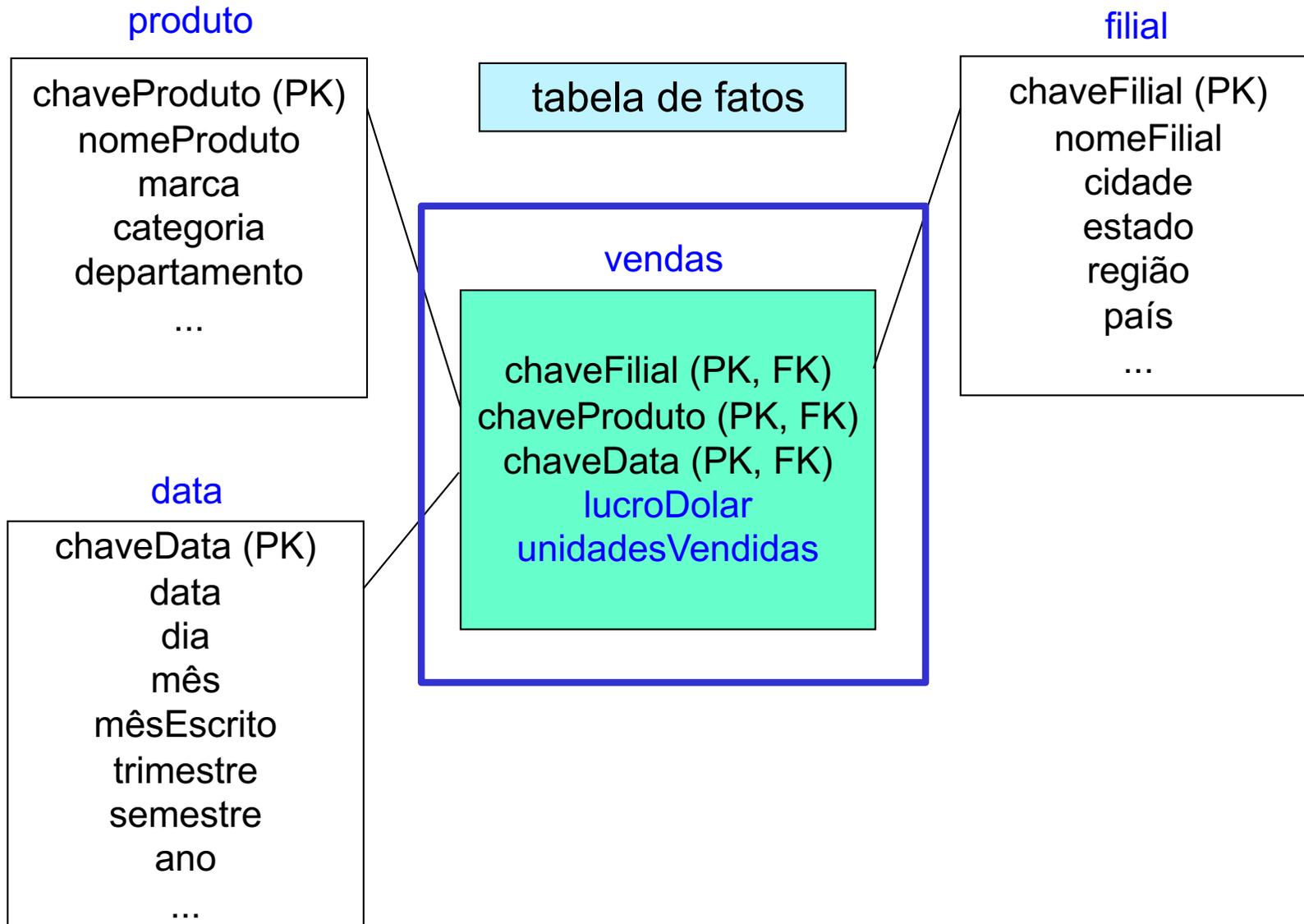
- Baseado no uso do modelo relacional
- Data warehouse
 - representado como uma coleção de esquemas de relação
- Tipos de esquema
 - **estrela** (star)
 - **floco de neve** (snowflake)
 - **estrela-floco** (starflake)

Tipos de Tabela
Tabela de Fatos
Tabela de Dimensão

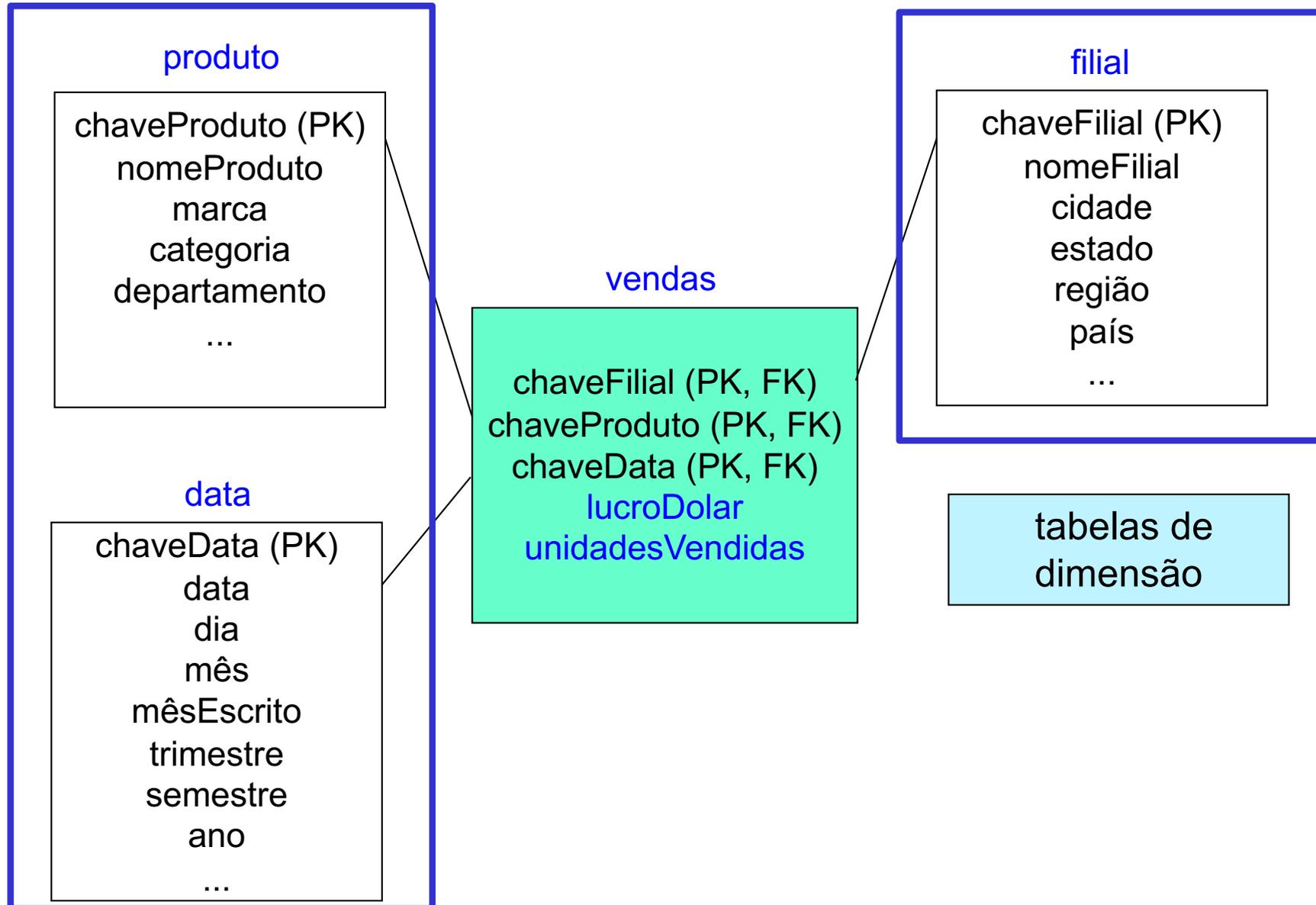
Esquema Estrela



Esquema Estrela



Esquema Estrela



Esquema Estrela

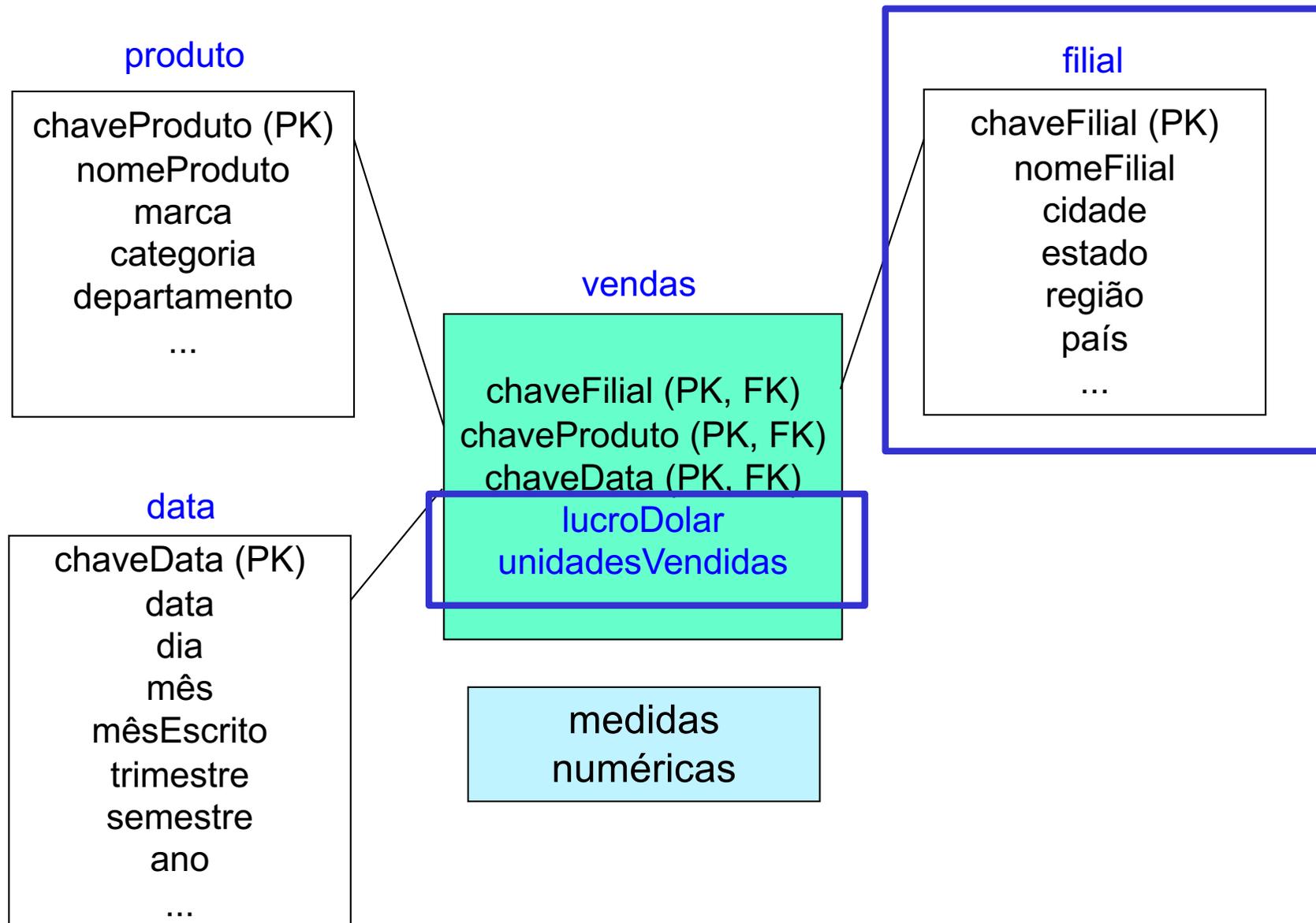


Tabela de Dimensão

- Localizada visualmente na **extremidade** da estrela
- Armazena
 - uma chave primária (chave artificial)
 - atributos da dimensão
- Características
 - usualmente larga e curta
 - com redundância
 - dados desmoralizados

Tabela de Dimensão **Filial**

chaveFilial	nomeFilial	cidade	estado	região	país	...
1	Vende Barato	São Carlos	São Paulo	Sudeste	Brasil	...
2	Para Poucos	São Carlos	São Paulo	Sudeste	Brasil	...
3	Atende Você	Araraquara	São Paulo	Sudeste	Brasil	...
4	Mega Vendão	Campinas	São Paulo	Sudeste	Brasil	...
5	Compra Fácil	Limeira	São Paulo	Sudeste	Brasil	...
6	Atende Você II	Araraquara	São Paulo	Sudeste	Brasil	...
7	Tudo Aqui	Belo Horizonte	Minas Gerais	Sudeste	Brasil	...
...

- Representa o aspecto geográfico (WHERE)

Tabela de Dimensão **Data**

chaveData	data	dia	mês	mêsEscrito	trimestre	semestre	ano	...
1	01/01/2020	1	1	janeiro	1	1	2020	...
2	02/01/2020	2	1	janeiro	1	1	2020	...
3	03/01/2020	3	1	janeiro	1	1	2020	...
4	04/01/2020	4	1	janeiro	1	1	2020	...
5	05/01/2020	5	1	janeiro	1	1	2020	...
6	06/01/2020	6	1	janeiro	1	1	2020	...
7	07/01/2020	7	1	janeiro	1	1	2020	...
...

- Representa o aspecto temporal (WHEN)
- Pode ser povoada antes das demais tabelas

Tabela de Dimensão **Produto**

chaveProduto	nomeProduto	marca	categoria	departamento	...
1	carne moída	Carnes e Cia	alimento	açougue	...
2	frango	Frangote	alimento	açougue	...
3	bolacha salgada	Bolachinhas	alimento	biscoitos	...
4	ovo	Ovo Dourado	alimento	hortifruti	...
5	bucha	Esfregando	limpeza	higiene do lar	...
6	sabão em pó	Bolhas e Bolhas	limpeza	higiene do lar	...
...

- Representa o aspecto sendo investigado (WHAT)
- Praticamente povoada com dados do ambiente operacional

Tabela de Fatos

- Localizada visualmente no **centro** da estrela
- Armazena
 - as medidas numéricas relevantes ao negócio (**fatos**)
 - uma chave estrangeira para cada tabela de dimensão
 - **chave primária: combinação das chaves estrangeiras**
- Características
 - usualmente fina e longa
 - sem redundância
 - sem dados esparsos

Tabela de Fatos **Vendas**

data	produto	filial	lucro	unidadesVendidas
1	1	1	1,5	5
1	2	1	2,1	6
1	3	1	1,7	7
2	1	1	1,5	5
2	2	1	2,1	5
2	3	1	1,7	7
3	1	2	2,3	4
3	2	1	2,1	5
3	3	1	1,7	7
4	1	2	2,3	4
4	2	1	2,1	5
4	3	2	2,9	10
...

Esquema Relacional

filial (chaveFilial, nomeFilial, cidade, estado, região, país, ...)

produto (chaveProduto, nomeProduto, marca, categoria, departamento, ...)

data (chaveData, dia, mês, mêsEscrito, trimestre, semestre, ano, ...)

vendas (chaveData, chaveProduto, chaveFilial,
lucroDolar, unidadesVendidas)

Dimensões Clássicas

- **When**

- ex.: tempo, data, hora

- **What**

- ex.: produto, funcionário, cliente

- **Where**

- ex.: filial, departamento, setor

- **Why**

- ex.: promoção, cargo

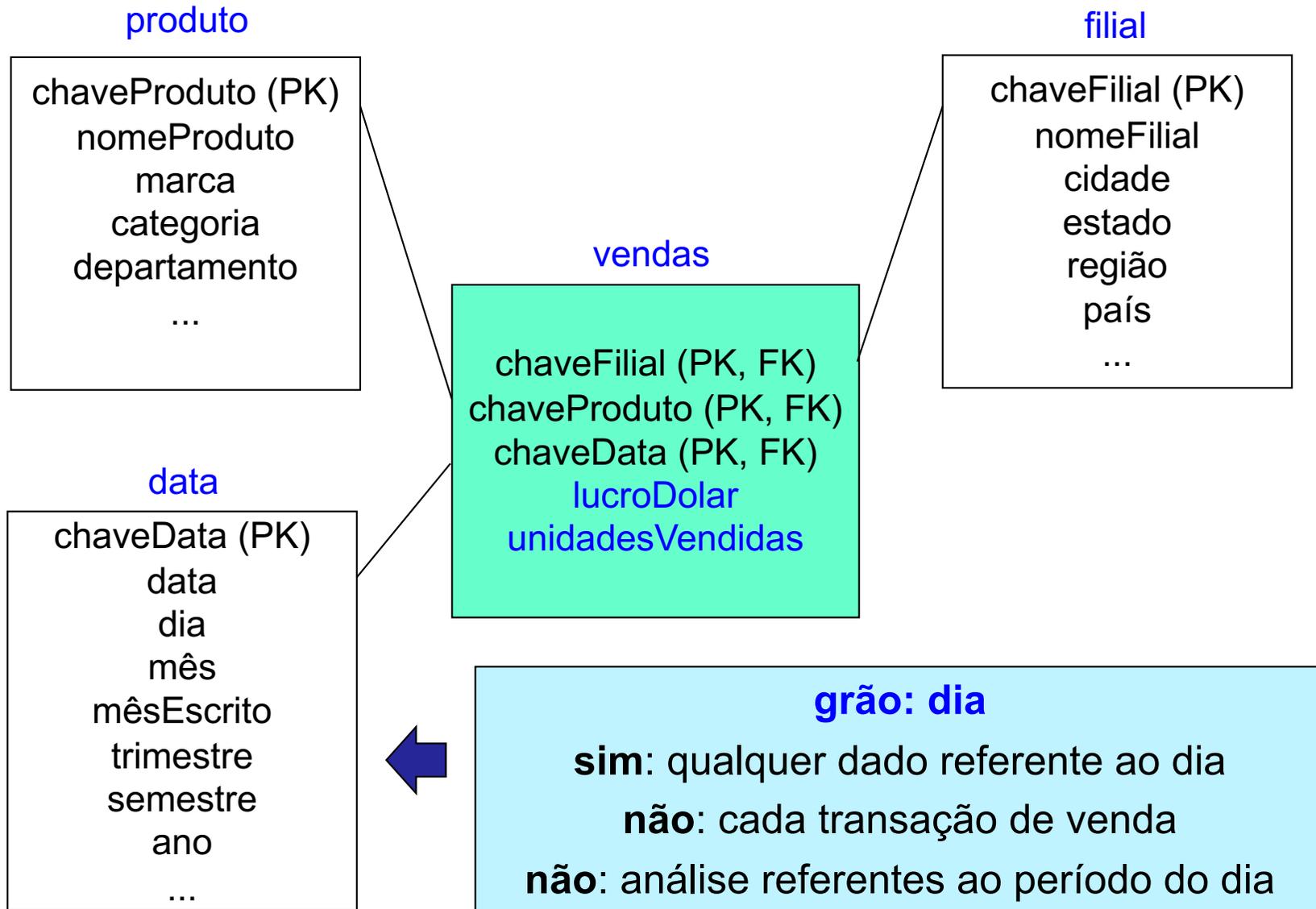
Granularidade

- Grau de detalhamento em que os dados são armazenados
 - Dados **mais detalhados** (ou **menos agregados**)
 - Dados **menos detalhados** (ou **mais agregados**)
- Aspecto de projeto muito importante
 - **Volume** de dados do *data warehouse*
 - **Consultas** que podem ser respondidas

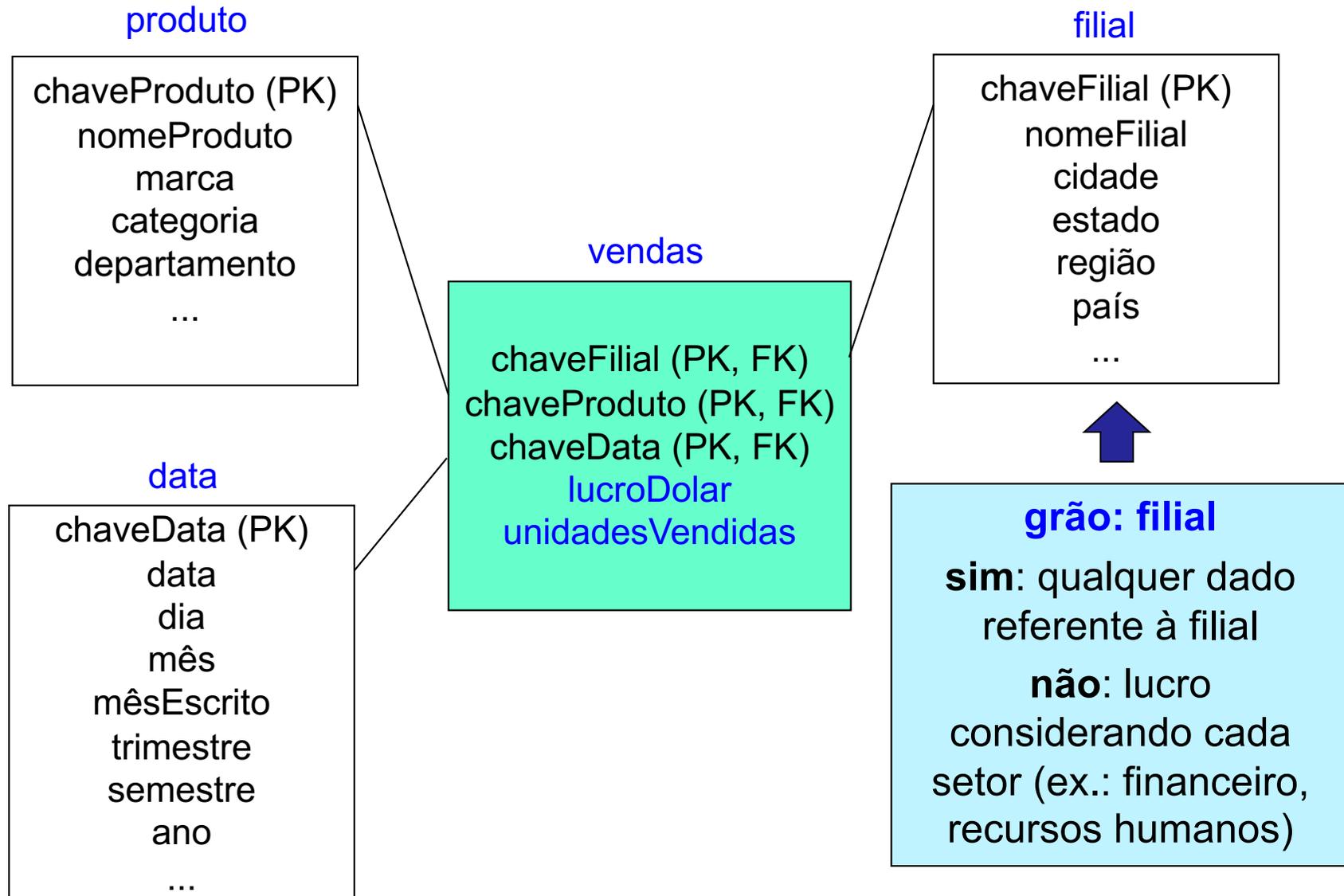
Tamanho do Grão

- Grão muito pequeno
 - Dados mais detalhados
 - *Data warehouse* é muito mais volumoso
 - Quantidade de consultas que podem ser respondidas é maior
- Grão muito grande
 - Dados menos detalhados
 - *Data warehouse* é menos volumoso
 - Quantidade de consultas que podem ser respondidas é menor

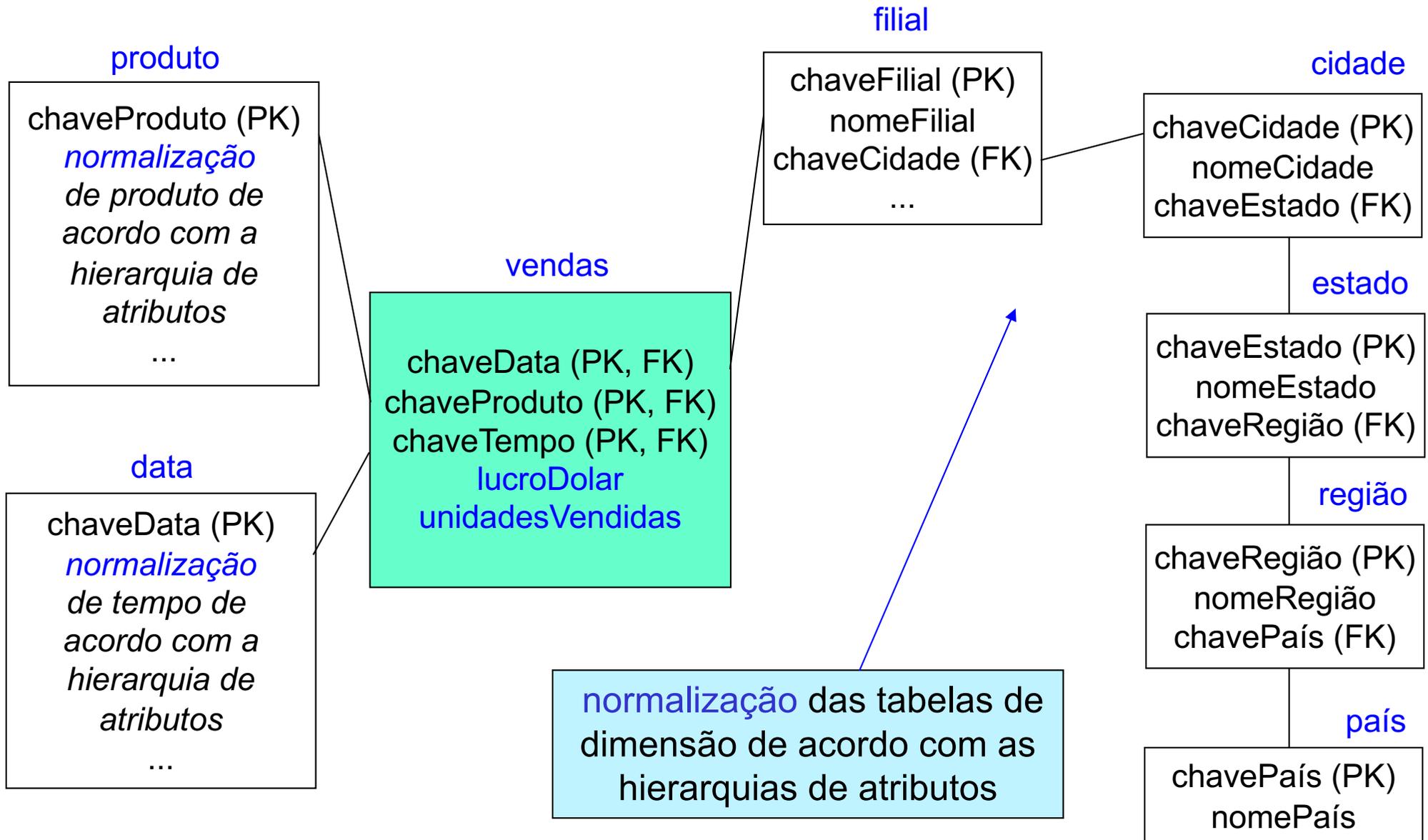
Grão da Tabela Data



Grão da Tabela Filial



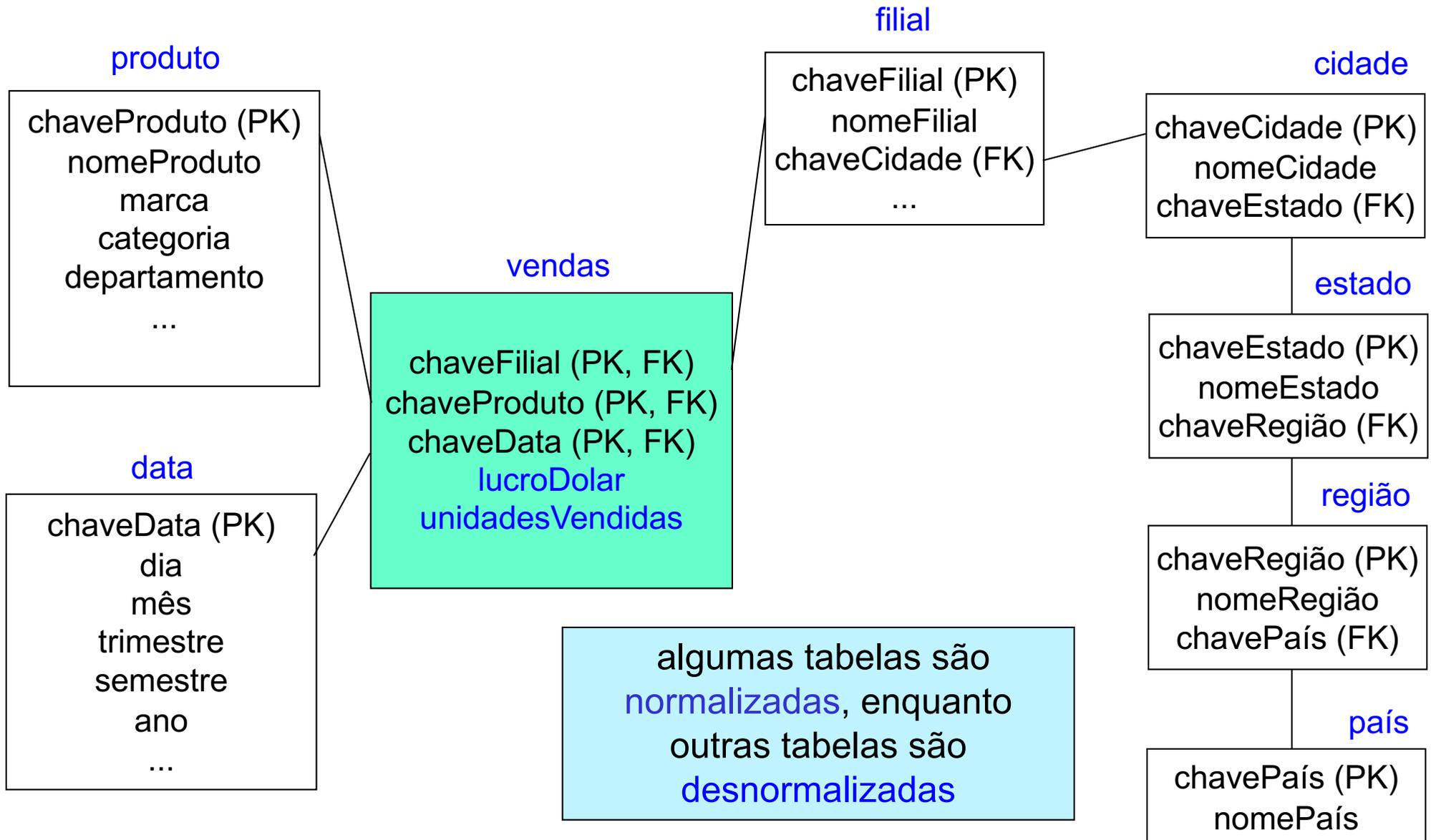
Esquema Floco de Neve



Estrela *versus* Floco de Neve

Característica	Estrela	Floco de Neve
redundância dos dados nas tabelas de dimensão	redundantes dados desnormalizados	sem redundância dados normalizados
processamento de junção-estrela	mais simples junção da(s) tabela(s) de dimensão necessária(s) com a tabela de fatos	mais complexa junção da(s) tabela(s) de dimensão necessária(s) com a tabela de fatos + junção da(s) tabela(s) de dimensão normalizada(s)

Esquema Estrela-Floco (Starflake)



Esquema Estrela-Floco

país (chavePaís, nomePaís)

região (chaveRegião, nomeRegião, chavePaís)

estado (chaveEstado, nomeEstado, chaveRegião)

cidade (chaveCidade, nomeCidade, chaveEstado)

filial (chaveFilial, nomeFilial, chaveCidade, ...)

produto (chaveProduto, nomeProduto, marca, categoria, departamento, ...)

data (chaveData, dia, mês, mêsEscrito, trimestre, semestre, ano, ...)

vendas (chaveData, chaveProduto, chaveFilial,
lucroDolar, unidadesVendidas)

Constelação de Fatos

