



FACTORY
MAKESMARTERDECISIONS
TRUSTED



ANÁLISE DE
DADOS, BUSINESS
INTELLIGENCE E IA

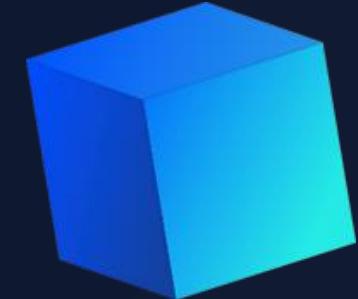


+9 anos

HÁ QUASE UMA
DÉCADA NO
MERCADO TECH



TRANSFORMAR
DADOS EM
INSIGHTS



FACTORY
MAKESMARTERDECISIONS
TRUSTED

TREINAMENTO
**FORMAÇÃO
DATA
EXPERTS**



André Rosa

Head of BI & Tech Learning na DataEX



Formação Acadêmica

- Graduado Sistema de informações – 2004-2007
- Pós Graduado em BI; 2008-2009
- Graduado em Gestão de TI; 2016-2019
- Pós Graduado em BI com Big data. 2018-2019

Experiências

- + de 20 anos de experiência Gestão e Consultoria em TI;
- ERP SAP e Totvs (Logix);
- + de 20 mil horas de projetos;
- Empresas nacionais e multinacionais de diversos portes e segmentos;
- +81K alunos.



Conhecendo Futuros Data Experts



Nome



Curso



Cidade



Trabalhando SIM/NÃO

Quizzes

Kahoot!



A qualquer momento.



QR Code.



Acesso pelo smartphone.

Agenda Formação Data Experts

Semana 1

DATA	INÍCIO	FIM	ASSUNTO	HABILIDADES
15/01/2024	19:30	21:30	Abertura	Introdução Analytics
16/01/2024	19:30	21:30	Modelagem de dados	Modelagem Dimensional
17/01/2024	19:30	21:30	Power BI	Power Query & DAX
18/02/2024	19:30	21:30	Power BI	DAX
19/02/2024	19:30	21:30	Power BI	Visuals & Deploy

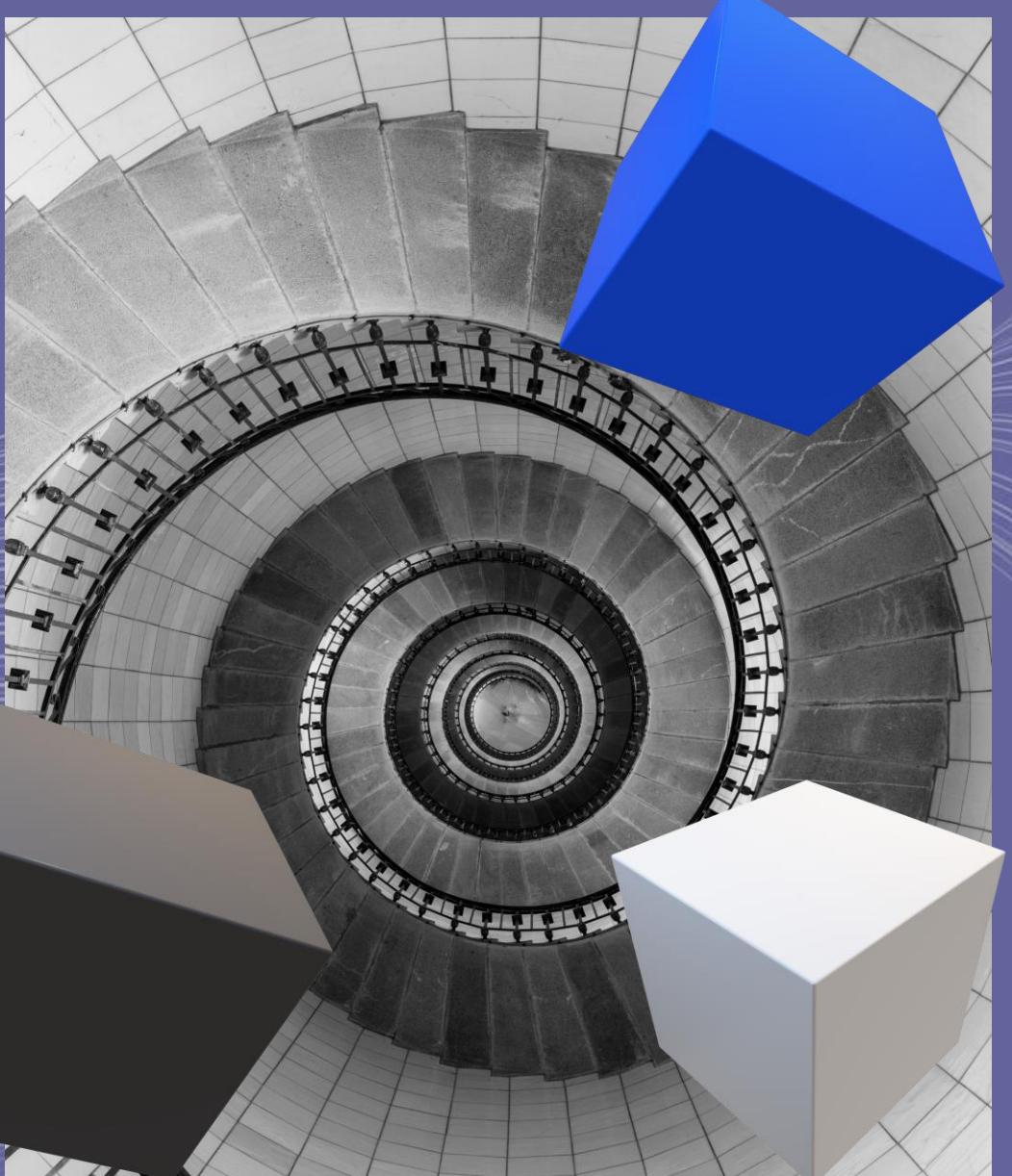
MODELAGEM DIMENSIONAL



3 Fases

- 🚀 Levantamento de necessidades.
- 🚀 Identificação de fontes
- 🚀 Modelagem dimensional





1. Entendimento da Necessidade

- O que será Analisado? (Fato)
- Como será Analisado? (Dimensões)
- Quais são os indicadores? (KPI's)

2. Mapeamento dos dados

- Identificar fontes/Origens
- Identificar Atributos/Colunas
- Gerar Matriz de dados/Dicionário de dados

3. Modelagem Dimensional

- **DW**
 - Gerar Star Schema ou Snow Flake.
 - Implementar SCD1 ou SCD2
- **Datalake**
 - Lakehouse
 - Medallion



Perguntas Objetivas

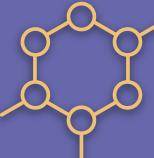
FASE 1

DIRECT QUESTIONS
(DQ)



1. Qual assunto será analisado? (Fato)

EX: Vendas



2. Como o assunto será analisado? (Dimensões)

EX: Clientes, Vendedores, Cidade, Tempo , produtos, categoria etc



3. Quais serão os indicadores? (KPI's)

EX: Total Vendas Valor, Total Vendas quantidade, Total descontos, Total Lucro, Total Custos etc...



4. Quais perguntas sobre os dados devem ser respondidas?

EX Total de vendas por vendedor por estado?

Total de vendas por vendedor por produto?

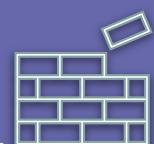
Variação do total de vendas Mês a Mês?



5. Critérios de segurança sobre o consumo de dados (RLS, DDM e CLS)

EX Regra: Vendedor visualiza apenas suas vendas.

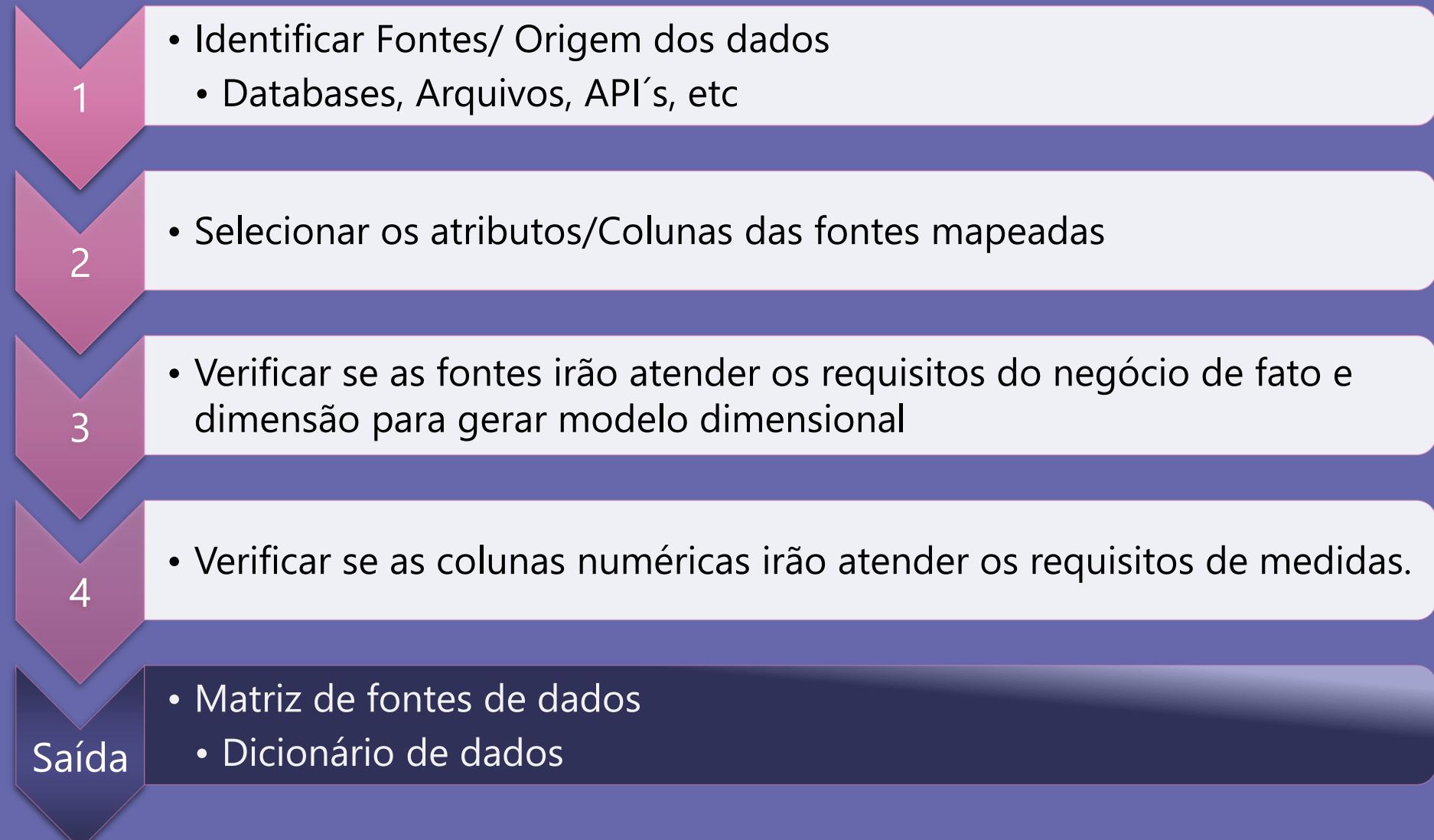
Gerente visualiza vendas apenas de vendedores de sua gestão;



Qual principal problema a ser resolvido?

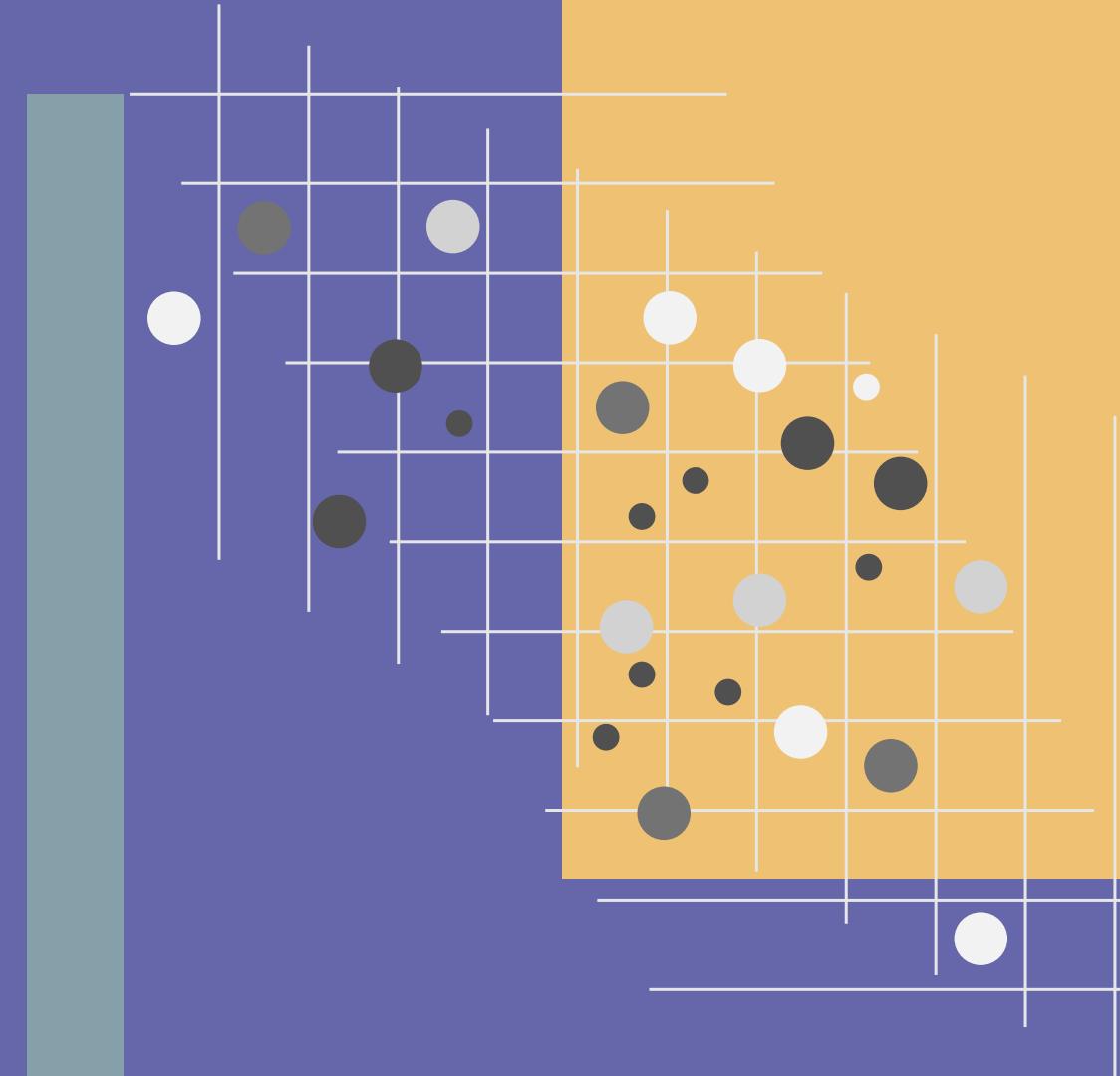
MAPEAMENTOS DE FONTES/ORIGENS

FASE 2



FASE 3

MODELAGEM DIMENSIONAL

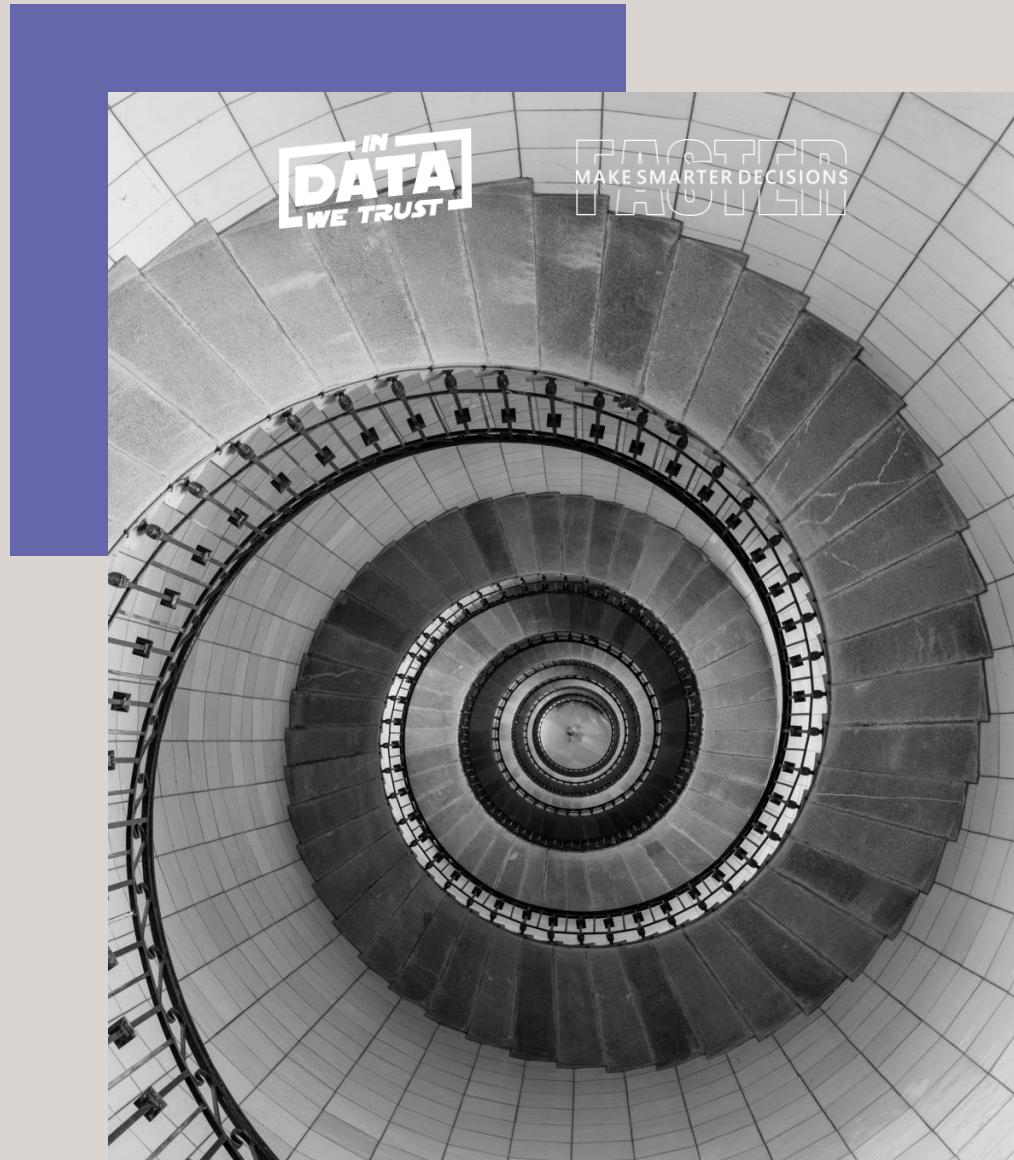


MODELAGEM DIMENSIONAL

Modelagem dimensional é uma técnica de **concepção e visualização** utilizada em Business Intelligence (BI) para **organizar dados em um modelo lógico** e facilitar a análise.

Tem como objetivo descrever aspectos comuns do de negócio.

Nesse tipo de modelagem, os dados **(Fatos)** são organizados em torno de **um conjunto de dimensões** (por exemplo, tempo, produto, cliente) e medidas (por exemplo, vendas, receita, lucro) que descrevem o negócio.



MODELAGEM DIMENSIONAL

O objetivo da modelagem dimensional é **fornecer uma estrutura de dados intuitiva e flexível** que possa ser facilmente explorada pelos usuários finais, permitindo que eles realizem análises e tomem decisões com base nas informações disponíveis. Essa abordagem é amplamente utilizada em data warehouses e outras soluções de BI para suportar a geração de relatórios, análises ad hoc e ferramentas de visualização de dados.

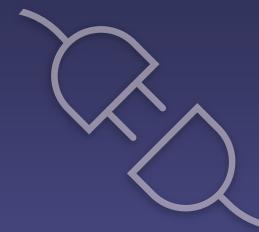


MODELAGEM DIMENSIONAL

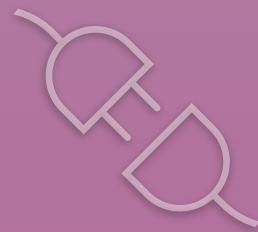


DATA
IN
WE TRUST
MADE SMARTER DECISIONS
PROTOTYPING

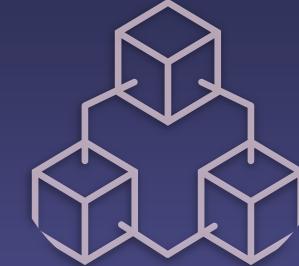
FASE 3



Etapa 3, depende
fortemente dos insumos
das Etapa 1 e 2.



Etapa 3, deve atender e
suportar a todos requisitos
da Etapa 1 e 2.



Saída

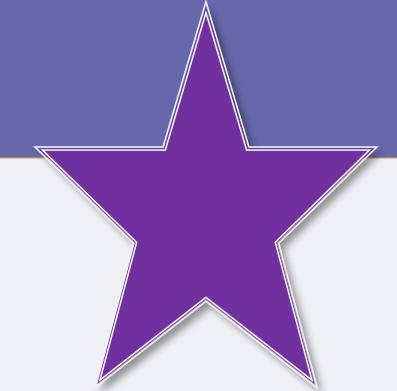
- MODELAGEM DIMENSIONAL LÓGICA/FÍSICA
 - STAR SCHEMA OU SNOW FLAKE
 - OU
 - MODELAGEM PARA DATALAKE
 - LINHAGEM DE DADOS

FASE 3

TIPOS DE MODELAGEM DIMENSIONAL (DW)



STAR SCHEMA



SNOW FLAKE



MODELAGEM CONCEITUAL

- A modelagem conceitual é a primeira etapa do processo e envolve a criação de um modelo abstrato que representa as principais entidades e relacionamentos do negócio. É uma visão geral dos dados que serão coletados e analisados pelo sistema de BI.

MODELAGEM LÓGICA

- A modelagem lógica é a segunda etapa e envolve a transformação do modelo conceitual em um modelo lógico mais detalhado, que define os dados de forma mais precisa e completa. O modelo lógico é independente de tecnologia e fornece uma estrutura conceitual para o sistema de BI.

MODELAGEM FÍSICA

- A modelagem física é a terceira etapa e envolve a implementação do modelo lógico em um ambiente de banco de dados específico. O modelo físico inclui detalhes como tipos de dados, chaves primárias, Surrogate Keys, index e constraints (FK*).
- *Avaliar a implementação de Foreign Keys (FK)

RESUMO

- Em resumo, a modelagem conceitual fornece uma visão geral do negócio, a modelagem lógica define a estrutura dos dados e a modelagem física implementa essa estrutura em um banco de dados específico. Juntas, essas etapas ajudam a garantir que o sistema de BI seja eficiente, preciso e atenda às necessidades dos usuários de negócios.

TIPO DE MODELOS



Conceitual



Vendas



Produtos

Lógico

Vendas

- Nota_fiscal
- Data
- Cod Produto
- Valor



Produtos

- Cod_prod
- descricao
- categoria

Físico

Vendas

- Nota_fiscal int primary key, not null
- Data date not null
- Cod Produto int not null
- Valor decimal(10,2) not null



Produtos

- Cod_prod int primary key, not null
- Descricao varchar(50)
- Categoria varchar(50)

FASE 3

RESUMO

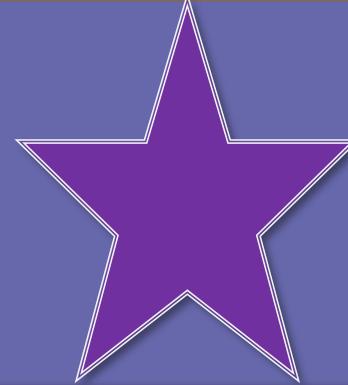


DATA
IN
WE TRUST
MADE SMARTER DECISIONS

Características	Conceitual	Lógica	Física
Nome da Entidade	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Relacionamentos de Entidade	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Atributos	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Chave primária		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Chave estrangeira		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Nome das tabelas			<input checked="" type="checkbox"/>
Nome das colunas			<input checked="" type="checkbox"/>
Tipo das colunas			<input checked="" type="checkbox"/>

FASE 3

STAR SCHEMA



STAR SCHEMA

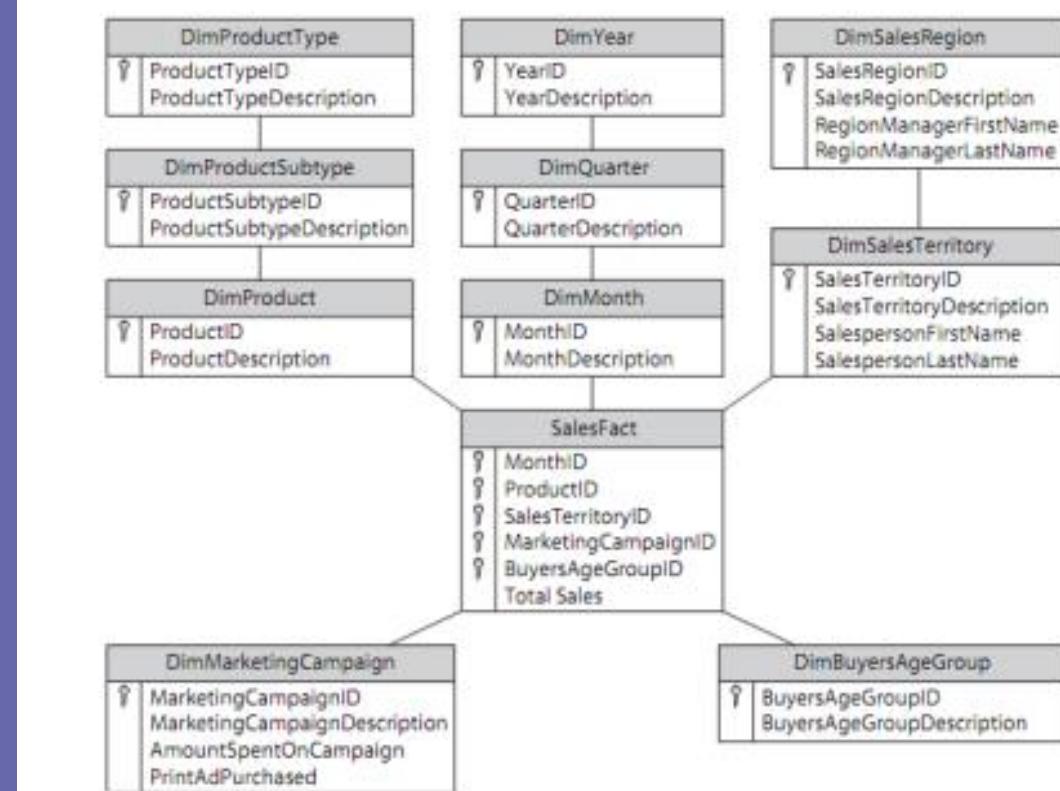
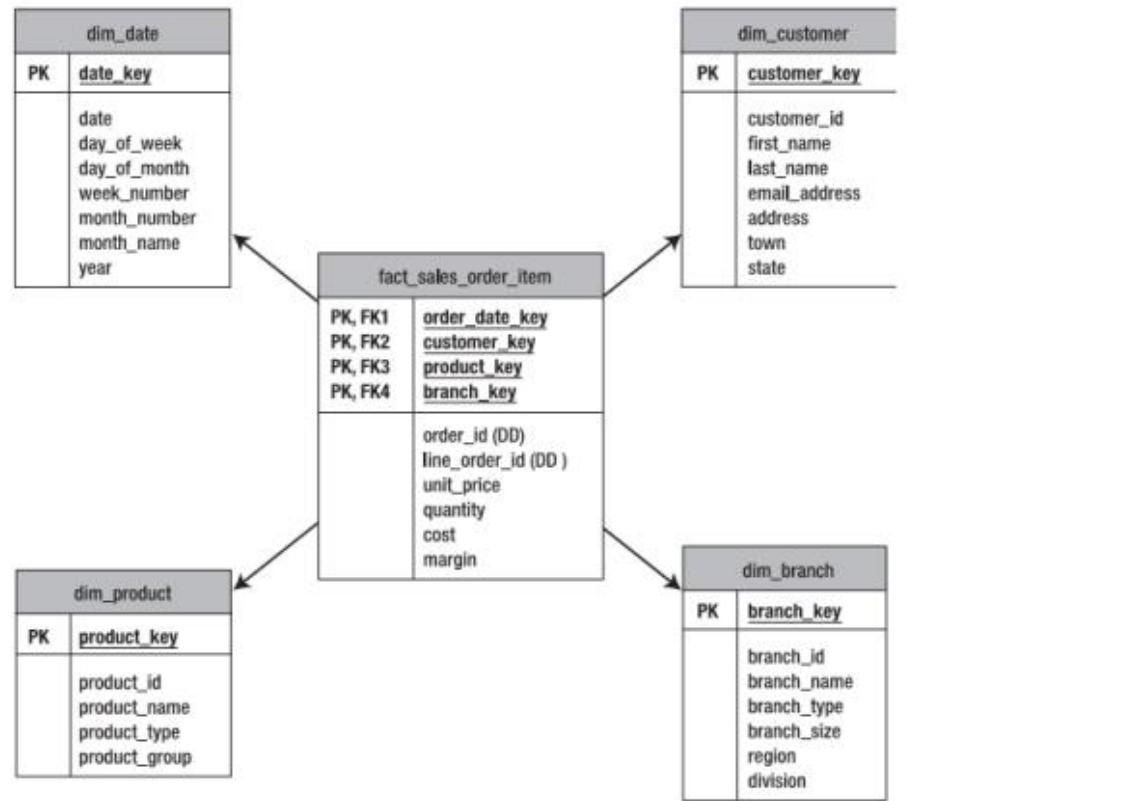
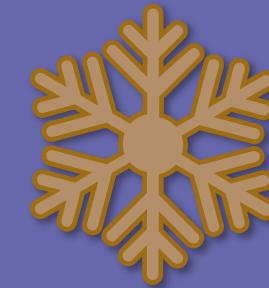
- É o modelo mais utilizado na **modelagem dimensional** para dar suporte à tomada de decisão e **melhorar a performance de sistemas voltados para consulta**.
- O esquema **estrela** é composto no centro por uma tabela **fato**, rodeada por tabelas de dimensão, ficando parecido com a forma de uma estrela

SNOW FLAKE



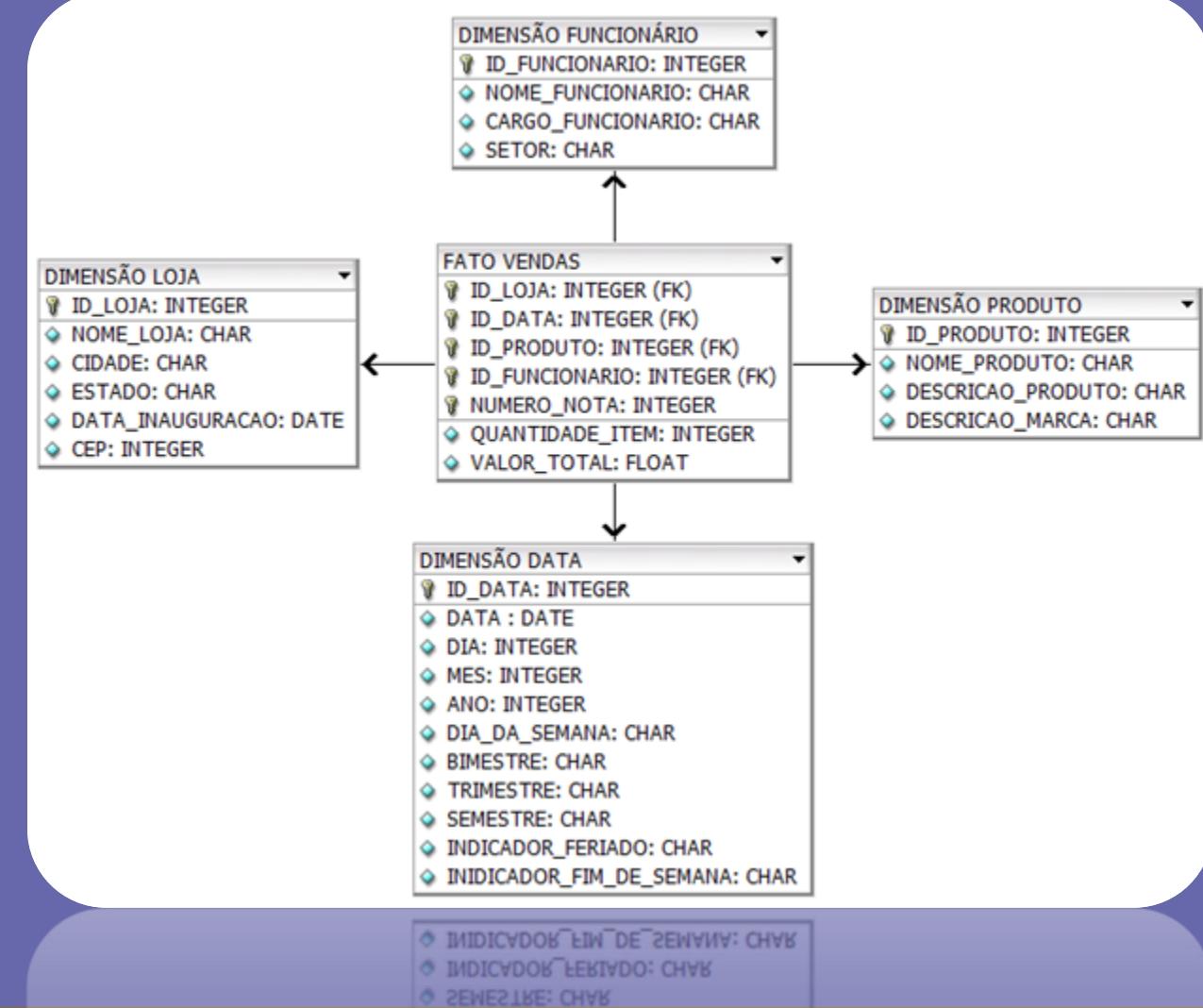
SNOW FLAKE

- Também é projetado para suportar tomada de decisão, mas economizando espaço em disco.
Para o Star Schema, o Snowflake é apenas mais um tipo de dimensão



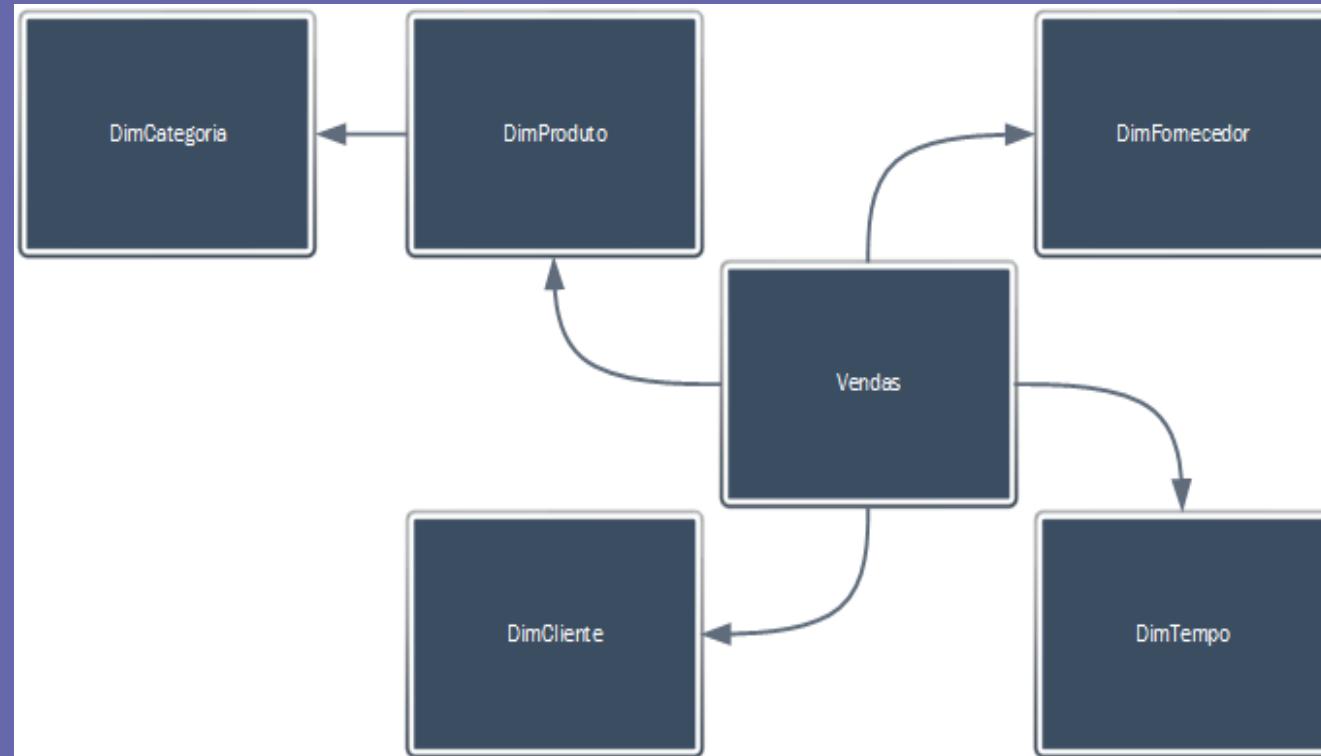


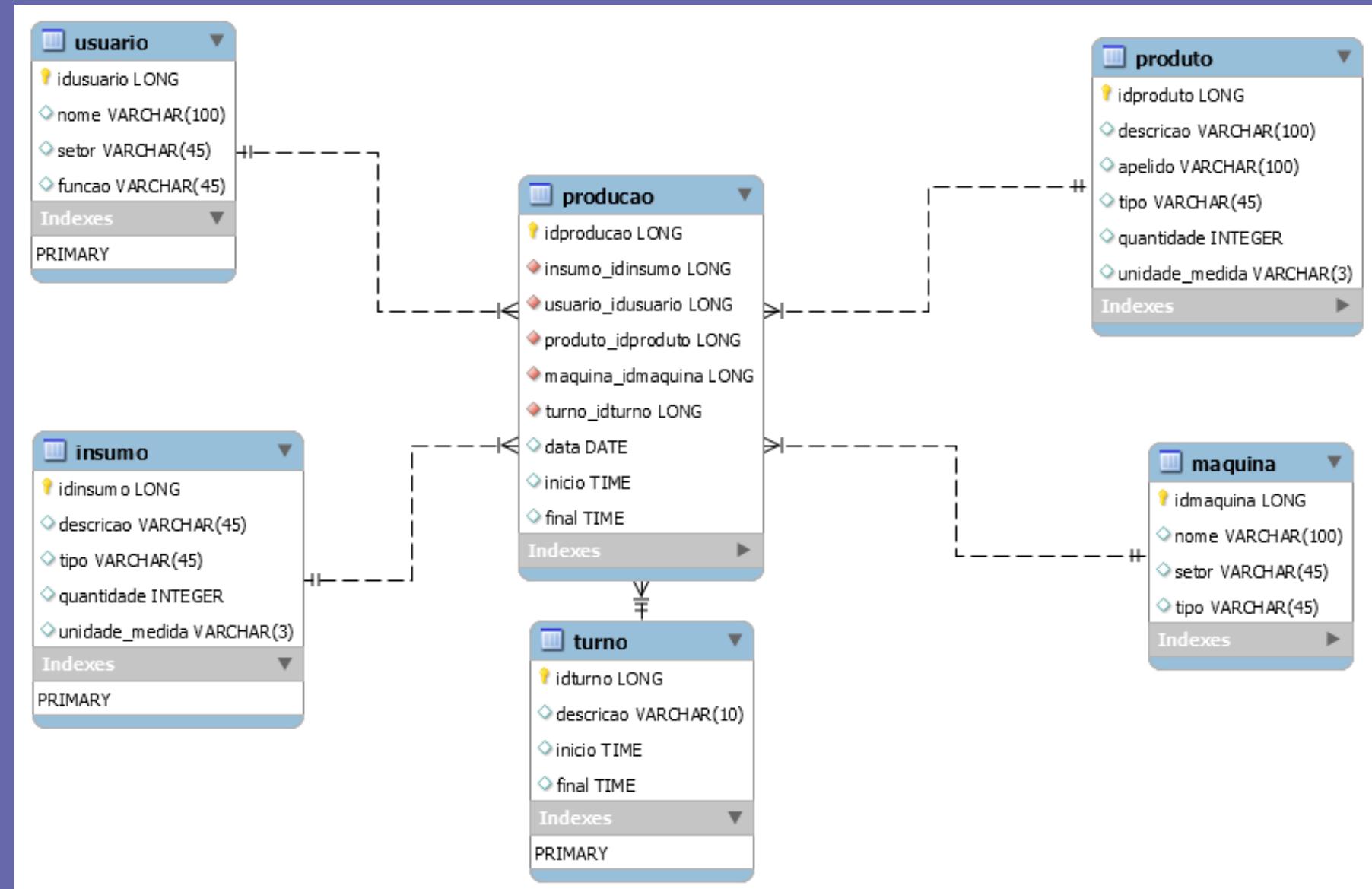
- Formato de uma estrela
- Tabela central é a Tabela Fato
- Fato serve de ligação entre as dimensões
- Dimensões se ligam unicamente à fato

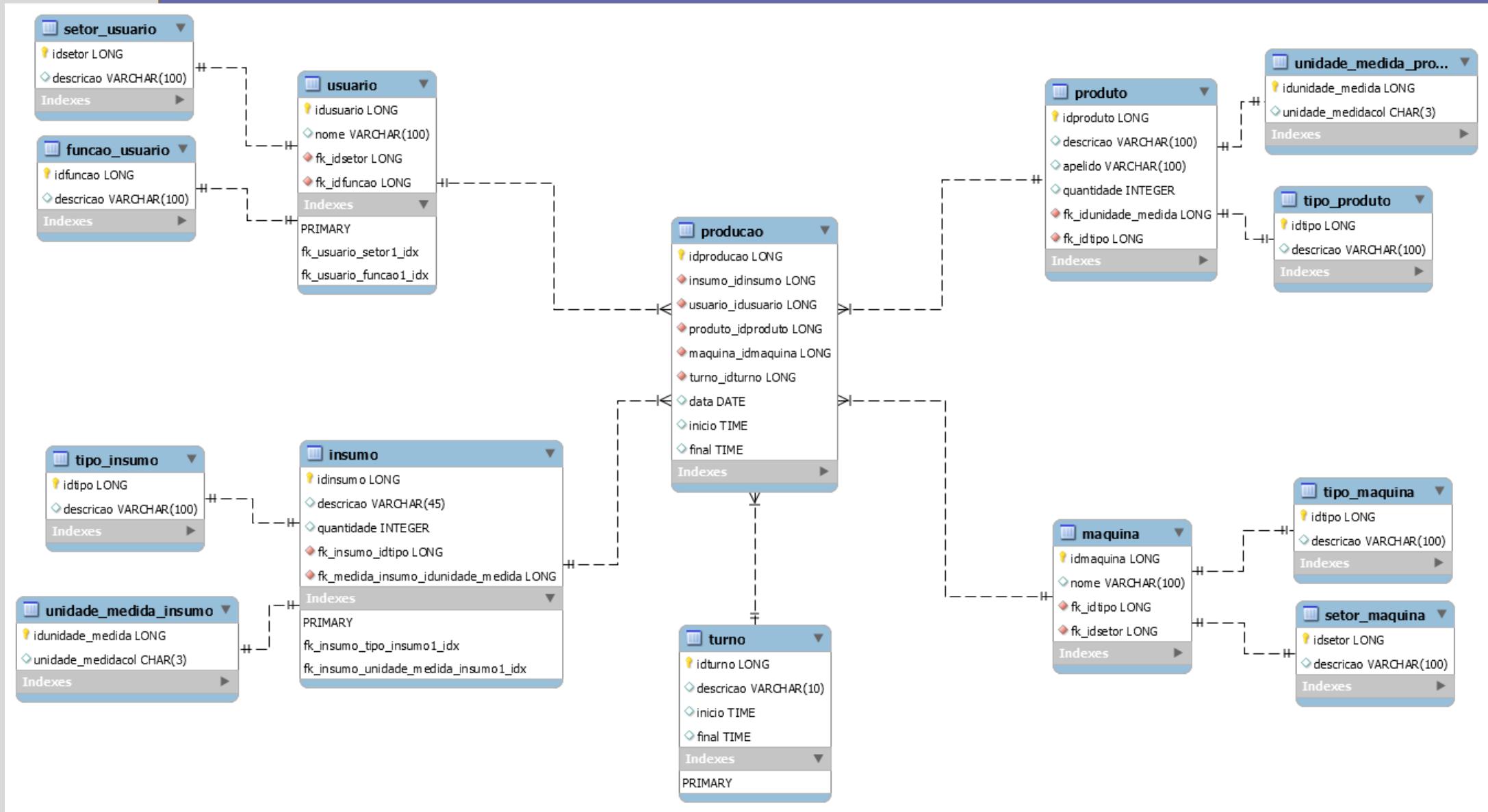




- Semelhante a um flocos de neve.
- Forma híbrida do “Star Schema”.
- Tabela fato não é a única tabela de ligação.
- Ligação entre dimensões pode levar a performance menor.
- Pode Otimizar o armazenamento.

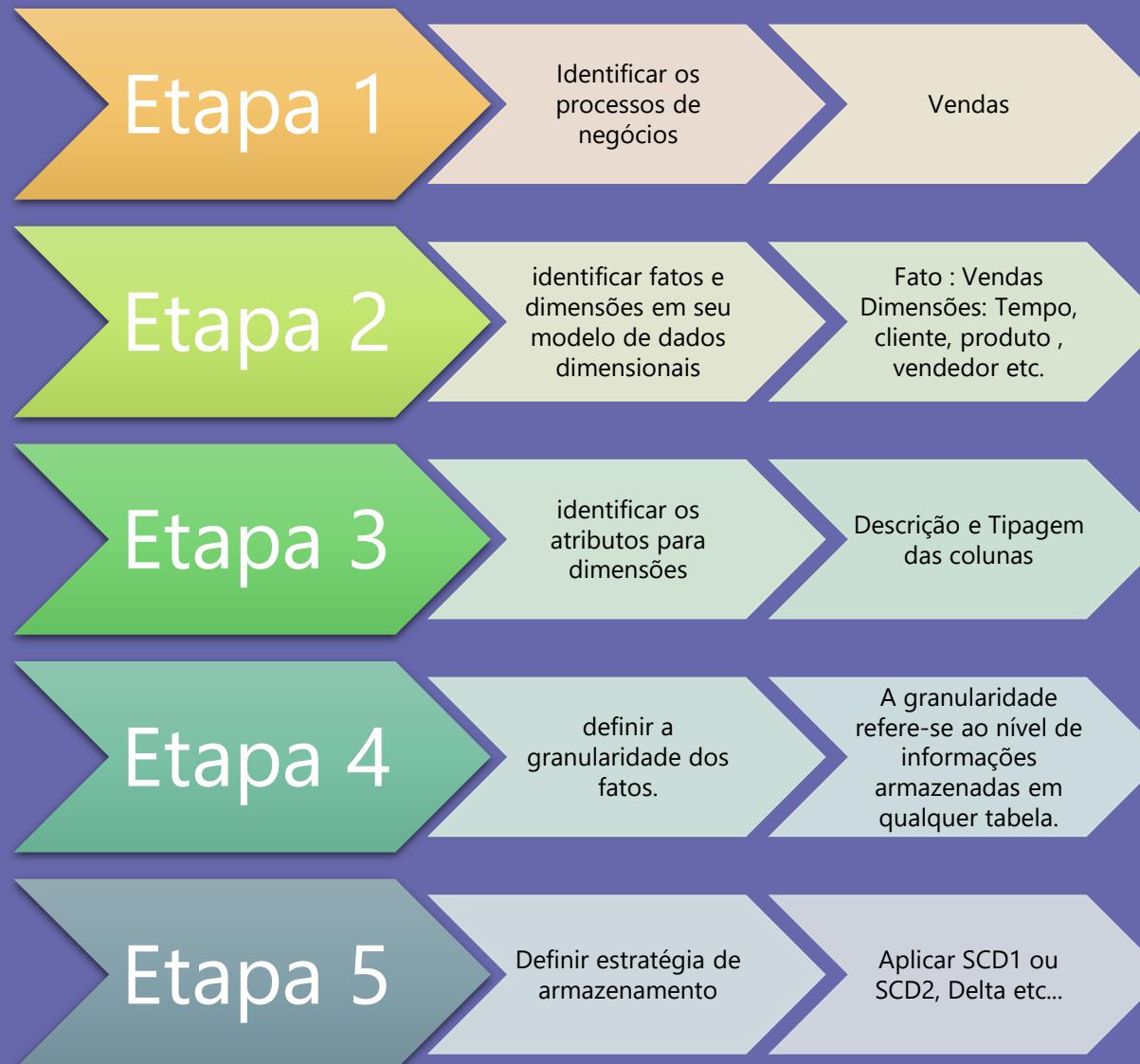






PROJETANDO UM MODELO DE DADOS DIMENSIONAL

FASE 3



DESNORMALIZAR “TABELÃO”



Desnormalizar e eliminar redundâncias de um controle de Vendas informal com informações muito relevantes sobre vendas de uma empresa global do segmento Ciclismo.

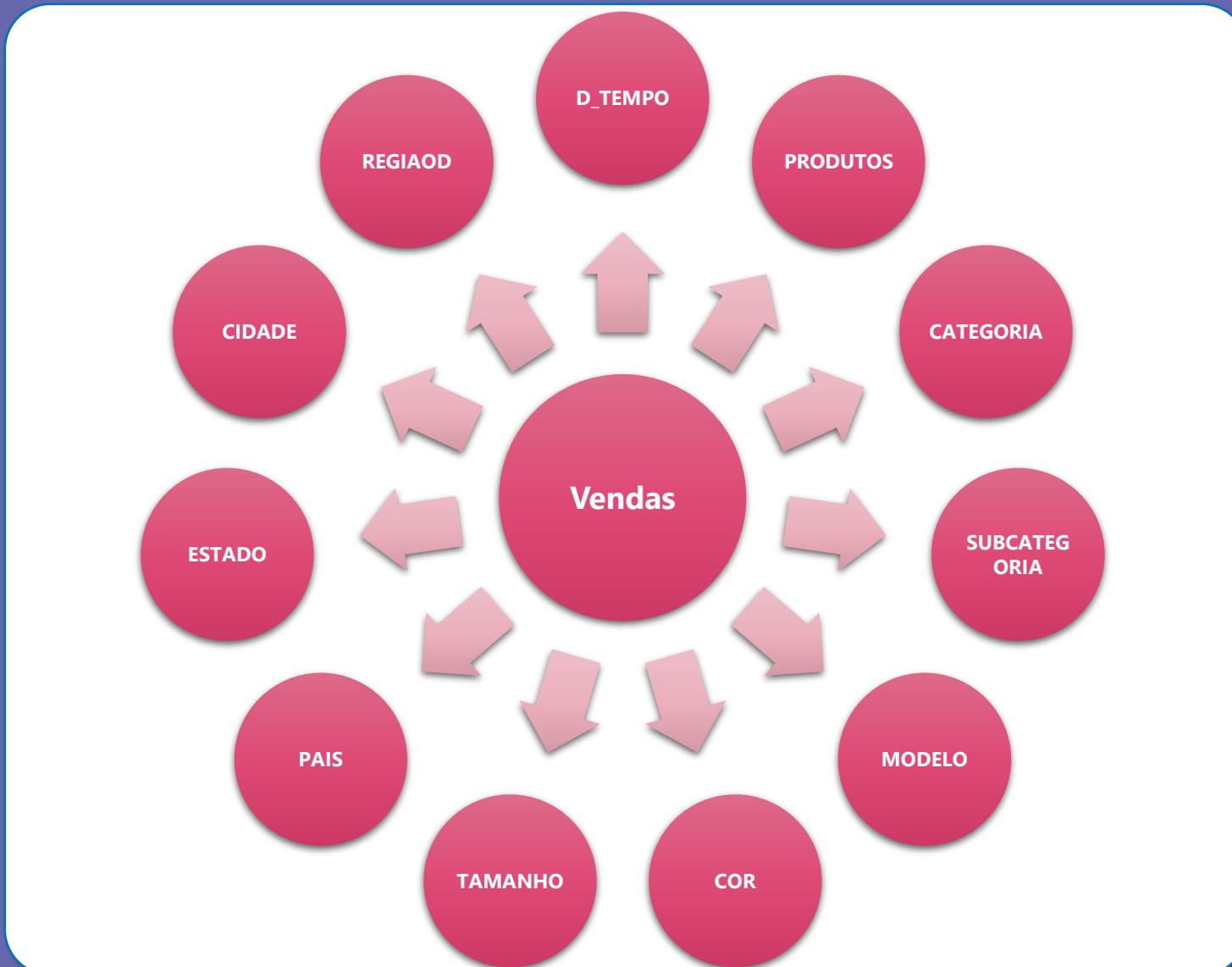
Propor modelo **conceitual**, físico e logico.

Categoria	SubCategoria	Modelo	Cor	Tamanho	Preço Custo Unit	Data Pedido	País	Estado	UF	Cidade	Região	Pedidos Qt	Total Venda
Bicicletas	Bicicletas de montanha	Mountain-100	Prata	48	1912,15	01/07/2017	Brasil	Paraíba	PB	Passagem	Nordeste	2	6799,98
Bicicletas	Bicicletas de estrada	Road-150	Vermelho	62	2171,29	01/07/2017	Brasil	Minas Gerais	MG	Córrego Fundo	Sudeste	1	3578,27
Bicicletas	Bicicletas de estrada	Road-150	Vermelho	52	2171,29	01/07/2017	Brasil	Alagoas	AL	Quebrangulo	Nordeste	2	7156,54
Bicicletas	Bicicletas de montanha	Mountain-100	Preto	38	1898,09	01/07/2017	Brasil	Ceará	CE	Varjota	Nordeste	1	3374,99
Bicicletas	Bicicletas de estrada	Road-150	Vermelho	48	2171,29	01/08/2017	Brasil	Rio Grande do No	RN	Jucurutu	Nordeste	1	3578,27
Bicicletas	Bicicletas de estrada	Road-150	Vermelho	56	2171,29	01/08/2017	Brasil	Minas Gerais	MG	Lambari	Sudeste	1	3578,27

IDENTIFICANDO AS ENTIDADES

MODELO CONCEITUAL

Categoria	Bicicletas
SubCategoria	Bicicletas de montanha
Modelo	Mountain-100
Cor	Prata
Tamanho	48
Preço Custo Unit	1912,15
Data Pedido	01/07/2017
País	Brasil
Estado	Paraíba
UF	PB
Cidade	Passagem
Região	Nordeste
Pedidos Qtd	2
Total Venda	6799,98



ANALYTICAL AREA

FASE 3

Fato

- Registra o acontecimento(**Fato**)
- Vendas, compras, acessos e etc....

Dimensão

- Ajuda **explicar** o acontecimento(**Fato**)
- Quando? Como? Quem? Onde?
- Quem vendeu? Qual cliente comprou?
- Tipo de produto?

SLOWLY CHANGING DIMENSIONS (SCD)



FASE 3

O SCD é uma sigla que significa Slowly Changing Dimensions (Dimensões que Mudam Lentamente, em português) e retrata as dimensões que sofrem atualizações em seus campos e os classifica pelo tipo de mudança existente em cada uma delas.

Vários tipos de SCD podem ser identificados no DW, variando de acordo com as características de atualizações das dimensões. As alternativas mais comuns de SCD são o SCD Tipo 1, SCD Tipo 2, SCD Tipo 3 e o SCD Híbrido.



FASE 3

SURROGATE KEYS (SK)



A Surrogate Key é uma chave artificial e auto incremental.

A palavra **artificial** vem do tipo, porque **ela não existe em lugar nenhum**, não está lá no transacional como a Natural Key, ela é criada no Data Warehouse.

E é auto incremental porque toda vez que é chamada, troca de número, então ela começa com 1 e vai indo para 2, 3, 4, e assim por diante.

tem as características de uma Primary Key.

é utilizada para referenciar a dimensão na fato

é auto incremental

é uma chave artificial

é criada no Data Warehouse

não pode se repetir

- O SCD Tipo 1 é a alteração que **não armazena histórico na dimensão**, ou seja, não é feito o versionamento do registro modificado. Trata-se do tipo mais simples, pois não há nenhum controle específico para a atualização dos dados, havendo apenas a sobreposição.



Tabela A		
ID	Nome	Salário
1	João	R\$ 1.800
2	Carlos	R\$ 2.500
3	Maria	R\$ 3.200

Tabela B		
ID	Nome	Salário
1	João	R\$ 1.800
2	Carlos	R\$ 2.500
3	Maria	R\$ 3.200

Tabela A		
ID	Nome	Salário
1	João	R\$ 2.200
2	Carlos	R\$ 2.500
3	Maria	R\$ 3.200

Tabela B		
ID	Nome	Salário
1	João	R\$ 2.200
2	Carlos	R\$ 2.500
3	Maria	R\$ 3.200

- O SCD Tipo 2 é a técnica mais utilizada para atualizações de dimensões. Nesse tipo de SCD é adicionado um novo registro com as mudanças, preservando sempre os dados anteriores. Dessa forma, os registros da tabela fato vão apontar para a versão correspondente nas dimensões de acordo com a data de referência.

Tabela A

ID	Nome	Salário
1	João	R\$ 1.800
2	Carlos	R\$ 2.500
3	Maria	R\$ 3.200

Tabela B

ID	Nome	Salário
1	João	R\$ 1.800
2	Carlos	R\$ 2.500
3	Maria	R\$ 3.200

Tabela A

ID	Nome	Salário
1	João	R\$ 2.200
2	Carlos	R\$ 2.500
3	Maria	R\$ 3.200

Tabela B

ID	Nome	Salário	Flag Atual
1	João	R\$ 1.800	0
1	João	R\$ 2.200	1
2	Carlos	R\$ 2.500	1
3	Maria	R\$ 3.200	1

1-) REGISTRO ORIGINAL (ERP)

CODIGO MATERIAL	MATERIAL
125	MONTAIN BIKE ARO 29

2-) REGISTRO NO DW APLICANDO SCD1

SK_MATERIAL	CODIGO MATERIAL	MATERIAL
001	125	MONTAIN BIKE ARO 29

3-) REGISTRO ORIGINAL ATUALIZADO (ERP)

CODIGO MATERIAL	MATERIAL
125	MTB ARO 29

4-) REGISTRO NO DW APLICANDO SCD1

SK_MATERIAL	CODIGO MATERIAL	MATERIAL
002	125	MTB ARO 29

1-) REGISTRO ORIGINAL (ERP)

CODIGO MATERIAL	MATERIAL
125	MONTAIN BIKE ARO 29

2-) REGISTRO NO DW APLICANDO SCD2

SK_MATERIAL	CODIGO MATERIAL	MATERIAL	DATA_INI	DATA_FIM	ATIVO
001	125	MONTAIN BIKE ARO 29	01/01/2023		1

3-) REGISTRO ORIGINAL ATUALIZADO (ERP)

CODIGO MATERIAL	MATERIAL
125	MTB ARO 29

4-) REGISTRO NO DW APLICANDO SCD2

SK_MATERIAL	CODIGO MATERIAL	MATERIAL	DATA_INI	DATA_FIM	ATIVO
001	125	MONTAIN BIKE ARO 29	01/01/2023	10/05/2023	0
002	125	MTB ARO 29	10/05/2023		1

EXEMPLOS DIMENSÕES



- ✓ Dimensão Tempo: que indica os períodos de tempo para a análise;
- ✓ Dimensão Produto: que indica quais produtos estão relacionados com as métricas;
- ✓ Dimensão Geografia: que indica a região geográfica onde se encontram as clientes estão.
- ✓ Dimensão por Canal de Negócio
- ✓ Dimensão por Linha de Produto
- ✓ Dimensão por Linha de Canal de vendas



EXEMPLOS DE MÉTRICAS/MEDIDAS



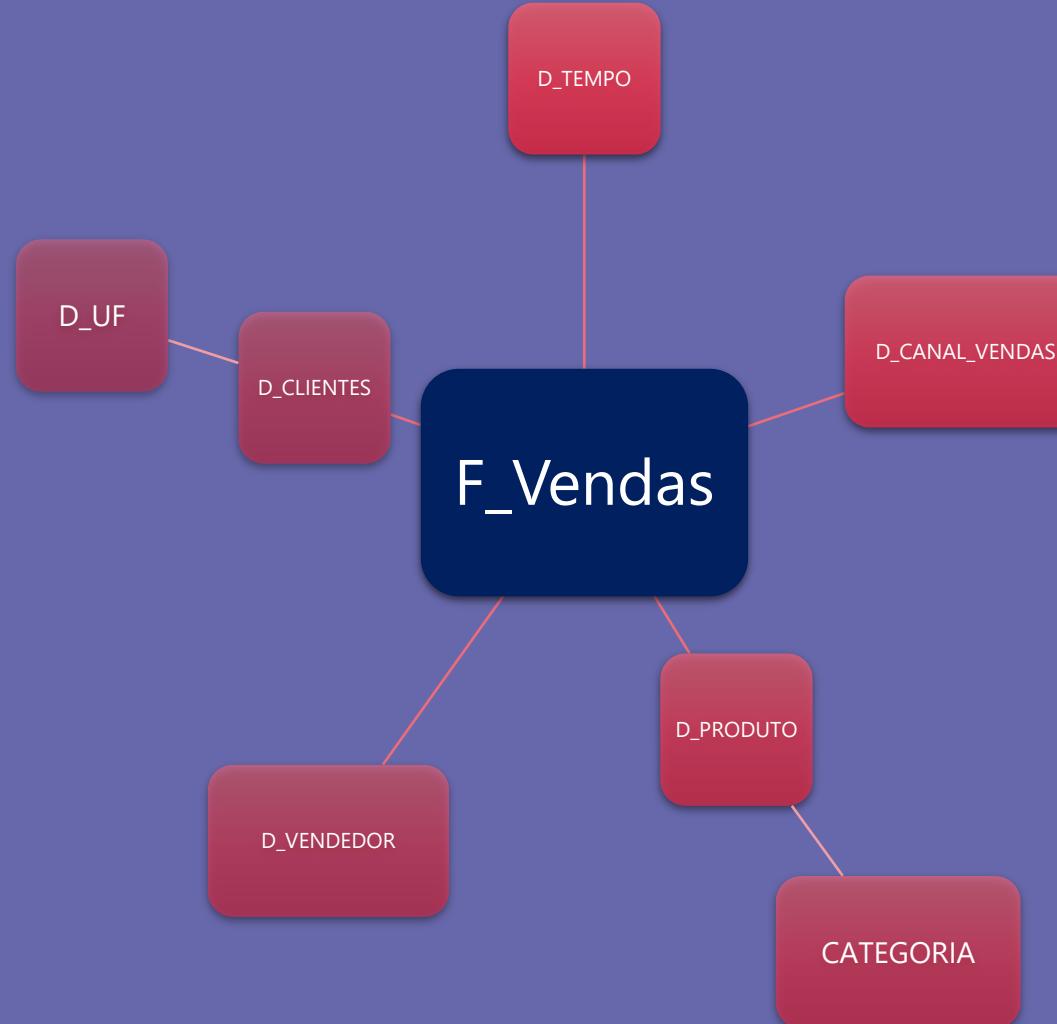
- ✓ Valor da venda – realizado;
- ✓ Valor da venda – previsto;
- ✓ Quantidade de itens – realizada;
- ✓ Quantidade de itens – prevista;
- ✓ Preço médio de venda;
- ✓ Custo médio;
- ✓ Margem de venda;
- ✓ % de variação entre o realizado e o planejado.
- ✓ % Variação Mês (MOM)
- ✓ %Variação Ano (YOY)
- ✓ Ticket Médio



MODELO DE DADOS (CONCEITUAL)



RAZER
MAKE SMARTER DECISIONS
RAZER



MODELO DE DADOS LÓGICO :TABELAS E SEUS ATRIBUTOS (SCD2)



PROGETTO
MAKE SMARTER DECISIONS

F_VENDAS	D_PRODUTOS	D_CATEGORIA	D_CLIENTES	D_UF	D_VENDEDOR	D_CANAL_VENDAS
<ul style="list-style-type: none"> •SK_prod •SK_Cliente •SK_Canal_vendas •SK_Vendedor •Pr_custo •Pr_Unit •Qtde •Dt_Venda •Situacao 	<ul style="list-style-type: none"> •SK_prod •SK_Categoria •Cod_prod •Nome_produto •Data_ini •Data_fim •Ativo 	<ul style="list-style-type: none"> •SK_Categoria •Cod_Categoria •Nome_Categoria •Data_ini •Data_fim •Ativo 	<ul style="list-style-type: none"> •SK_Cliente •SK_uf •Cod_Cliente •Nome_Cliente •Data_ini •Data_fim •Ativo 	<ul style="list-style-type: none"> •SK_uf •Cod_uf •Sigla •Nome_Estado •Data_ini •Data_fim •Ativo 	<ul style="list-style-type: none"> •SK_Vendedor •Cod_Vendedor •Nome_Vend •Data_ini •Data_fim •Ativo 	<ul style="list-style-type: none"> •SK_Canal_vendas •Cod_Canal_vendas •Desc_Canal_vendas •Data_ini •Data_fim •Ativo

Tabela Fato

- Contem SK's para relacionar com as **dimensões**
- **Não contem** os código do ERP/Originais
- Contêm as atributos/colunas para
 - Filtros
 - Cálculos

Dimensões

- Contem SK's para relacionar com a **fato**
- **Contem** os código do ERP/Originais
- Columnas descriptivas
- Atributos para implementar (SCD1 ou SCD2)

MODELO DE DADOS :TABELAS E SEUS ATRIBUTOS (SCD2)

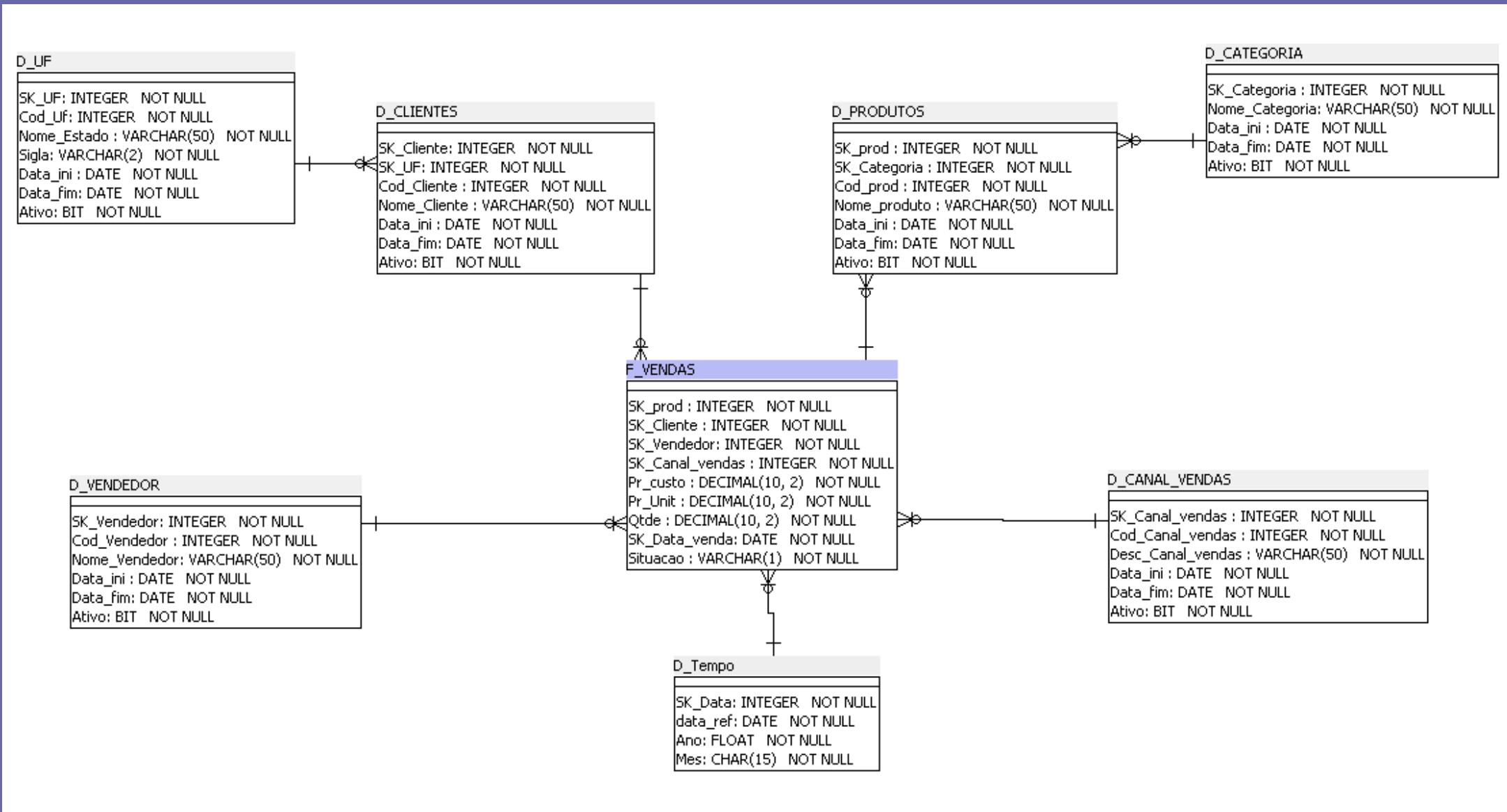


DIAGRAMA GERADO NO POWER ARCHITECT

MODELO DE DADOS FÍSICO: TABELAS E SEUS ATRIBUTOS

```
CREATE TABLE dbo.D_Tempo (
    SK_Data INT NOT NULL,
    data_ref DATETIME NOT NULL,
    Ano FLOAT NOT NULL,
    Mes CHAR(15) NOT NULL
)

CREATE TABLE dbo.D_CANAL_VENDAS_ (
    SK_Canal_vendas_ INT NOT NULL,
    Cod_Canal_vendas_ INT NOT NULL,
    Desc_Canal_vendas_ VARCHAR(50) NOT NULL,
    Data_ini_ DATETIME NOT NULL,
    Data_fim DATETIME NOT NULL,
    Ativo BIT NOT NULL
)

CREATE TABLE dbo.D_VENDEDOR_ (
    SK_Vendedor INT NOT NULL,
    Cod_Vendedor_ INT NOT NULL,
    Nome_Vendedor VARCHAR(50) NOT NULL,
    Data_ini_ DATETIME NOT NULL,
    Data_fim DATETIME NOT NULL,
    Ativo BIT NOT NULL
)

CREATE TABLE dbo.D_UF_ (
    SK_UF INT NOT NULL,
    Cod_Uf INT NOT NULL,
    Nome_Estado VARCHAR(50) NOT NULL,
    Sigla VARCHAR(2) NOT NULL,
    Data_ini_ DATETIME NOT NULL,
    Data_fim DATETIME NOT NULL,
    Ativo BIT NOT NULL
)
```

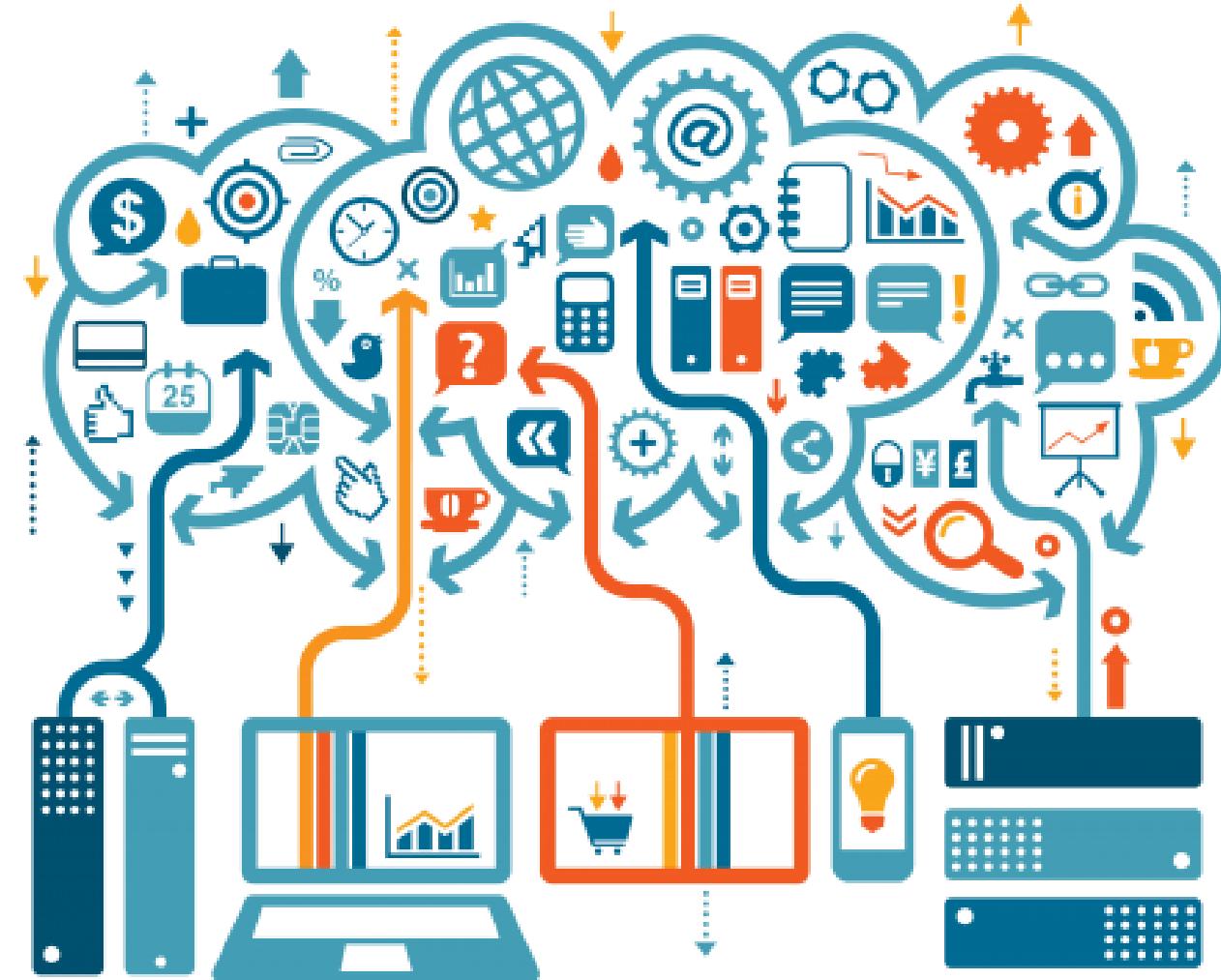
```
CREATE TABLE dbo.D_CLIENTES (
    SK_Cliente INT NOT NULL,
    SK_UF INT NOT NULL,
    Cod_Cliente_ INT NOT NULL,
    Nome_Cliente VARCHAR(50) NOT NULL,
    Data_ini_ DATETIME NOT NULL,
    Data_fim DATETIME NOT NULL,
    Ativo BIT NOT NULL
)

CREATE TABLE dbo.D_CATEGORIA_ (
    SK_Categoria_ INT NOT NULL,
    Nome_Categoria VARCHAR(50) NOT NULL,
    Data_ini_ DATETIME NOT NULL,
    Data_fim DATETIME NOT NULL,
    Ativo BIT NOT NULL
)

CREATE TABLE dbo.D_PRODUTOS_ (
    SK_prod_ INT NOT NULL,
    SK_Categoria_ INT NOT NULL,
    Cod_prod_ INT NOT NULL,
    Nome_produto_ VARCHAR(50) NOT NULL,
    Data_ini_ DATETIME NOT NULL,
    Data_fim DATETIME NOT NULL,
    Ativo BIT NOT NULL
)
```

```
CREATE TABLE dbo.F_VENDAS (
    SK_prod_ INT NOT NULL,
    SK_Cliente_ INT NOT NULL,
    SK_Vendedor INT NOT NULL,
    SK_Canal_vendas_ INT NOT NULL,
    Pr_custo_ DECIMAL(10,2) NOT NULL,
    Pr_Unit_ DECIMAL(10,2) NOT NULL,
    Qtde_ DECIMAL(10,2) NOT NULL,
    SK_Data_venda DATETIME NOT NULL,
    Situacao_ VARCHAR(1) NOT NULL
)
```

DÚVIDAS?



TREINAMENTO
**FORMAÇÃO
DATA
EXPERTS**

