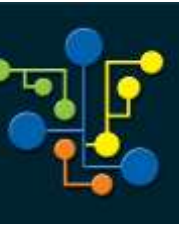




Data Science Academy



Design e Implementação de Data Warehouses



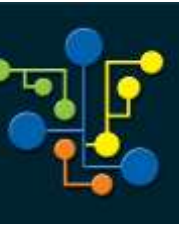
Business Intelligence e Modelagem de Dados





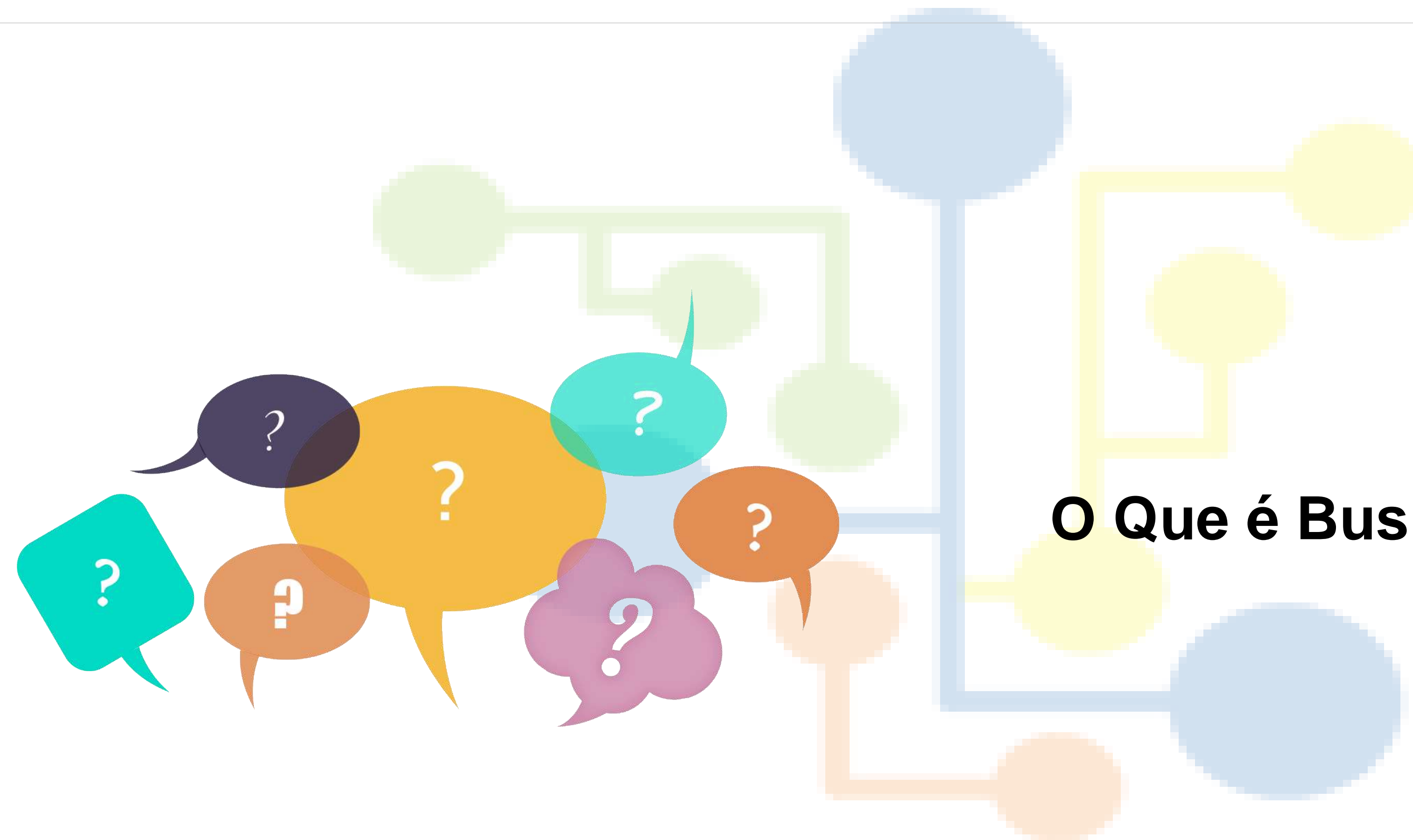
Definição de Conceitos Fundamentais



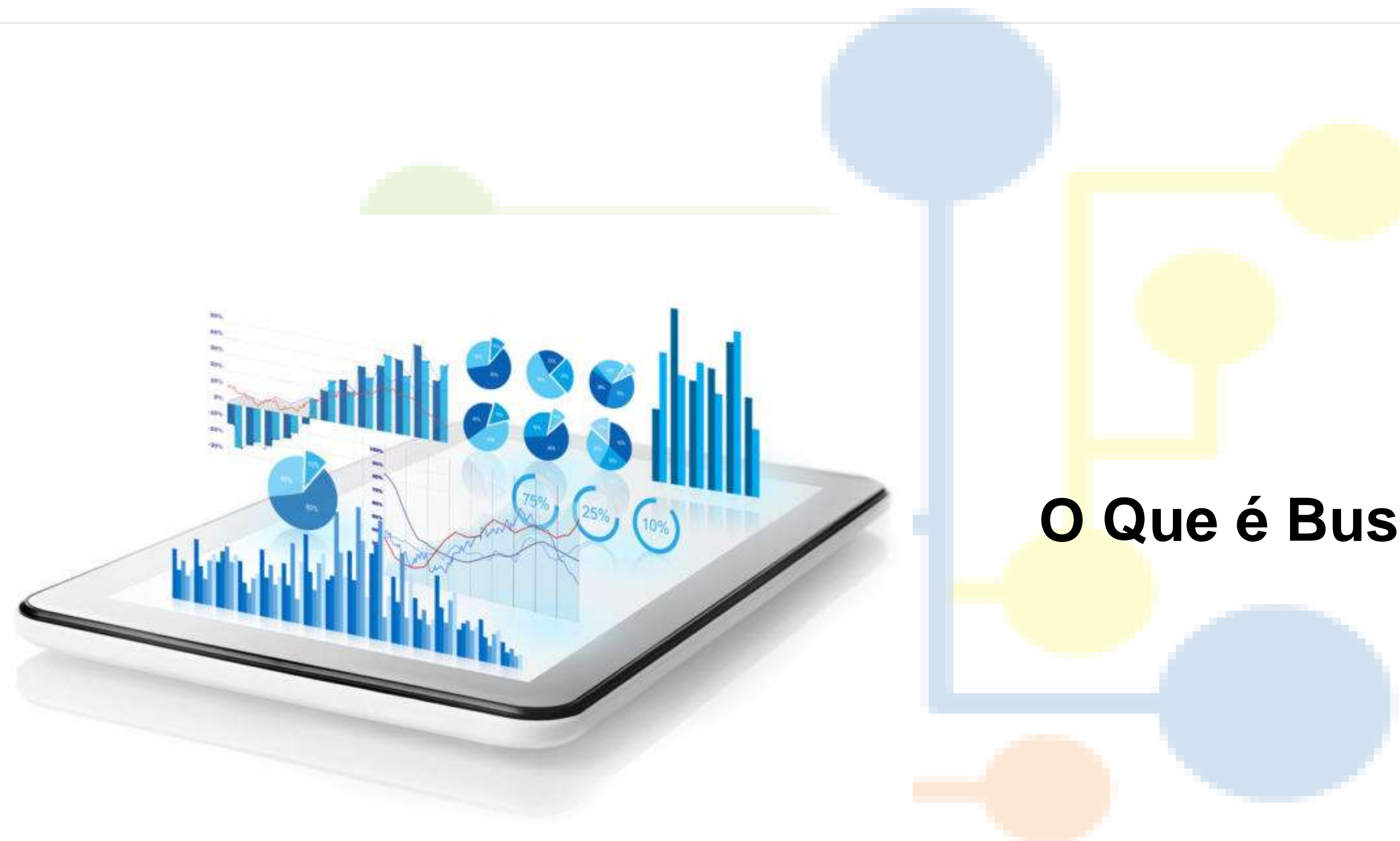


O Que é Business Intelligence?





O Que é Business Intelligence?



O Que é Business Intelligence?



O Que é Business Intelligence?

O Business Intelligence (BI) oferece o apoio nas decisões de forma inteligente por meio de um processo de captação de dados.

As informações são adquiridas de qualquer sistema e gravadas em um banco de dados modelado para o seu negócio.



O conceito de BI foi definido pelo Gartner Group como:

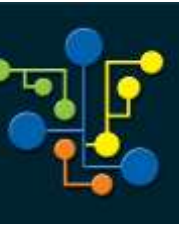
"Um termo genérico que inclui aplicações, infraestrutura, ferramentas e melhores práticas que permite o acesso e análise de informações para melhorar e otimizar decisões e desempenho".



“BI é um conjunto de técnicas e ferramentas que permite a organização e análise das informações para o suporte a tomada de decisão”.

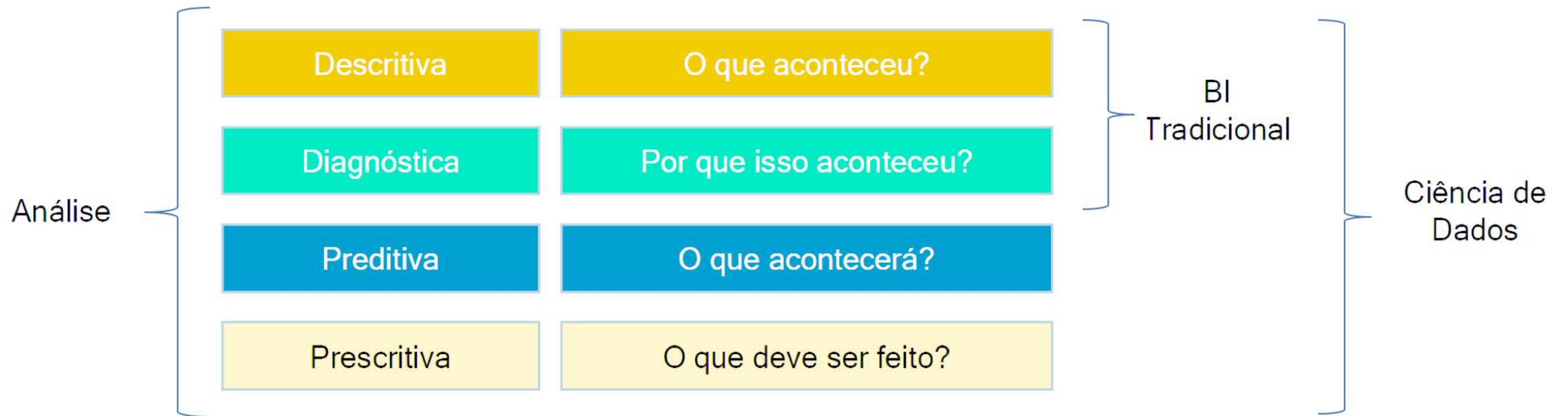


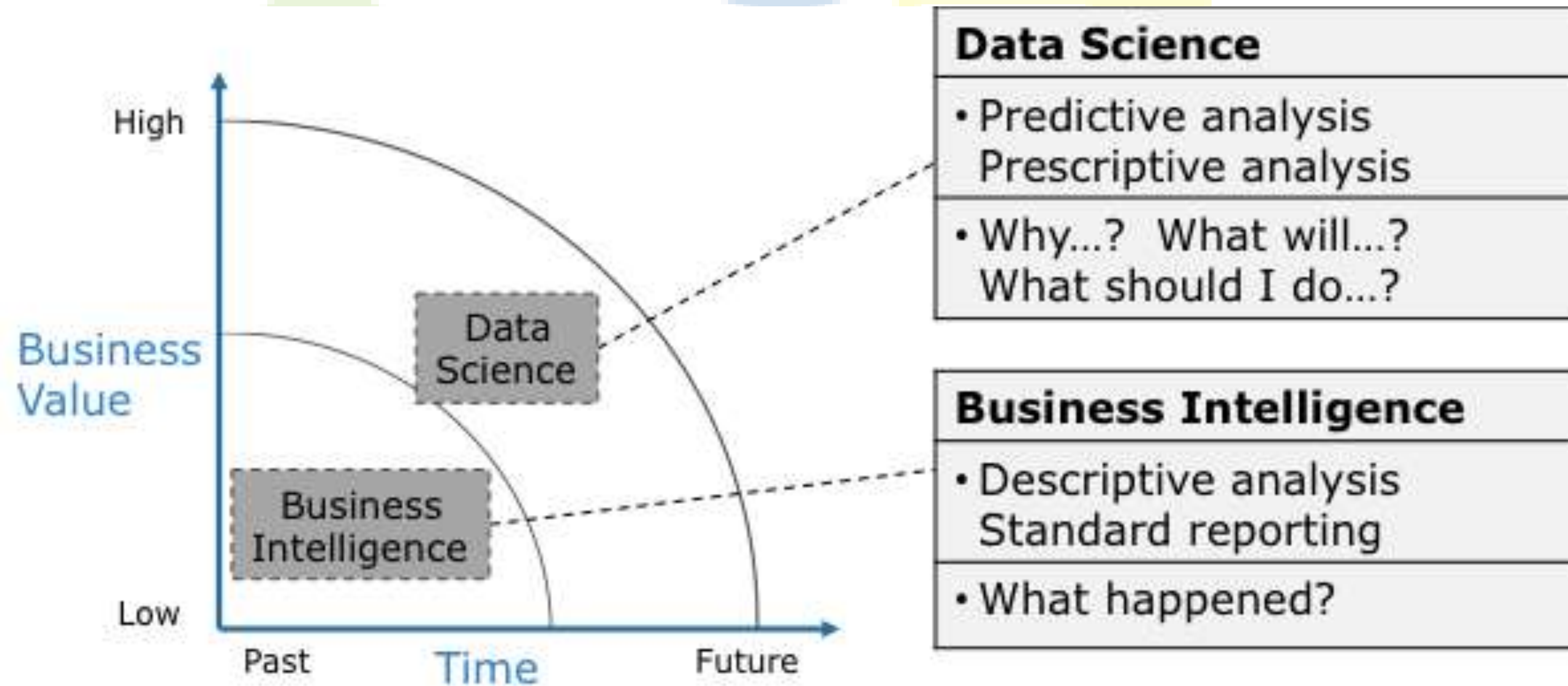


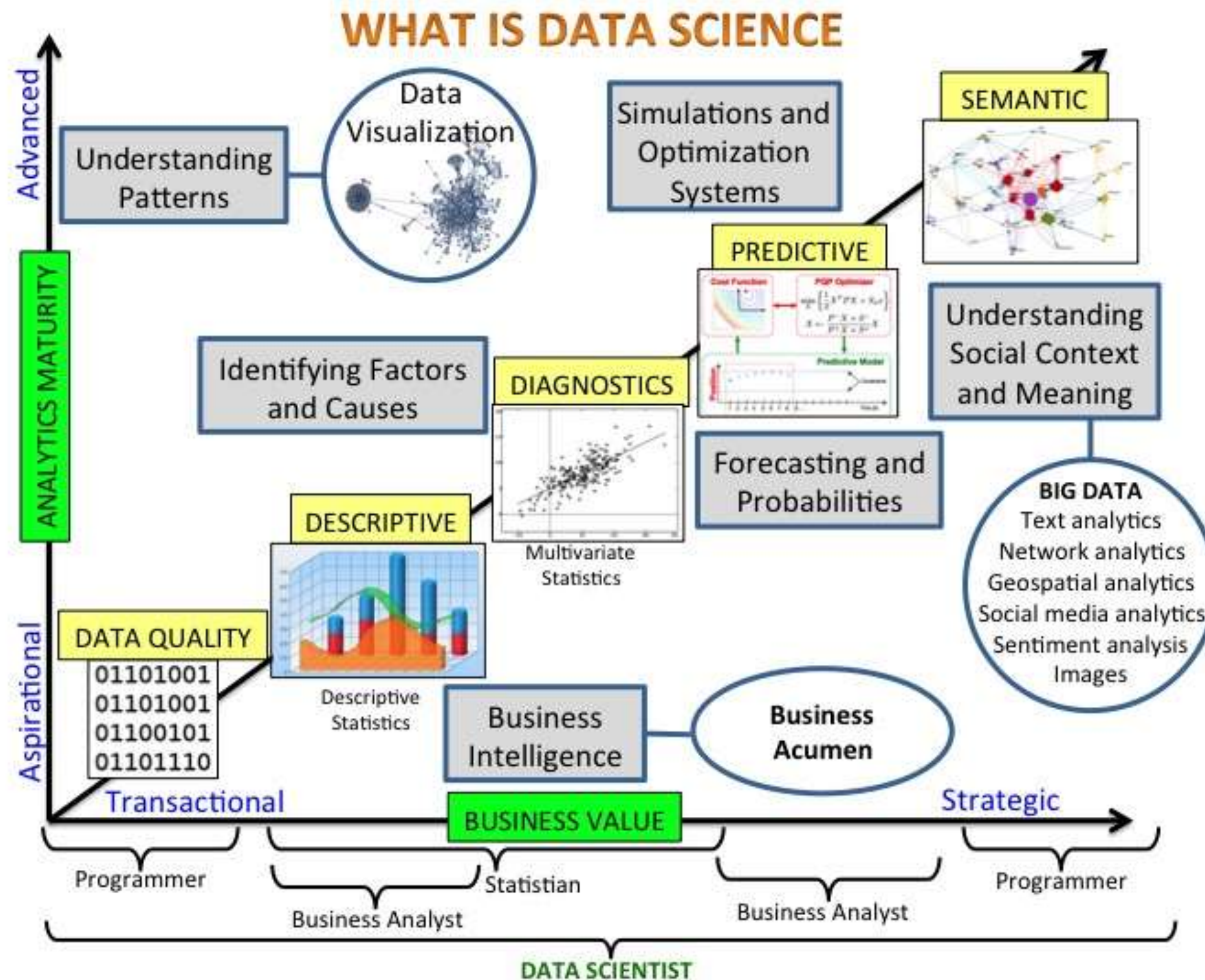


Business Intelligence X Data Science









O trabalho do Engenheiro de Dados é ajudar o Cientista de Dados a fazer o que está neste gráfico ao lado!



Ferramentas de BI X Ferramentas de Data Science



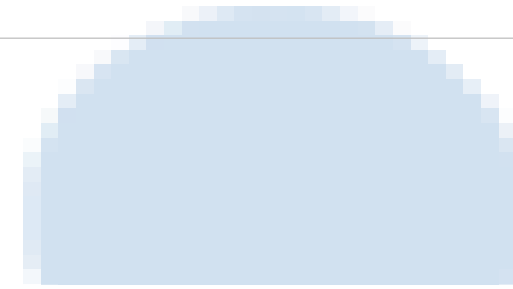


Visualização de Dados, Relatórios e BI



Data Science, Machine Learning e IA





Visualização de Dados, Relatórios e BI

Data Science, Machine Learning e IA

Data Warehouse

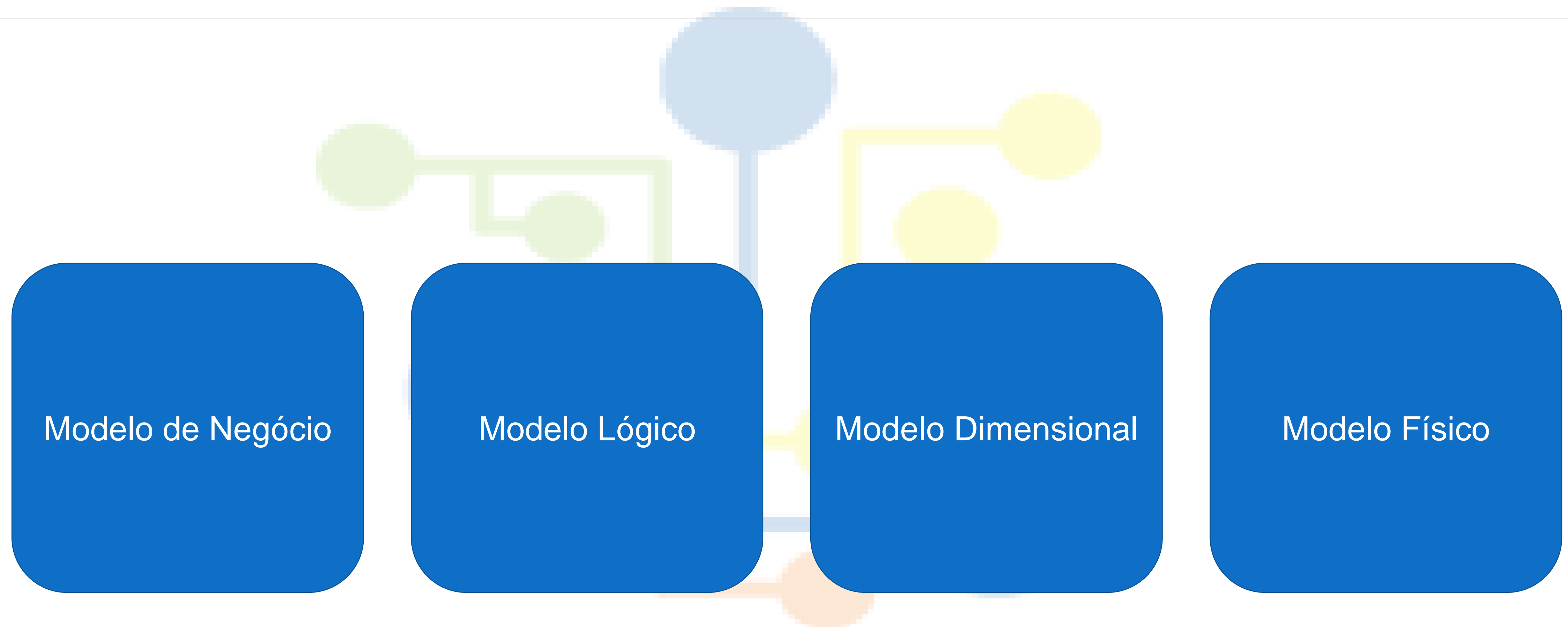
Data Lake

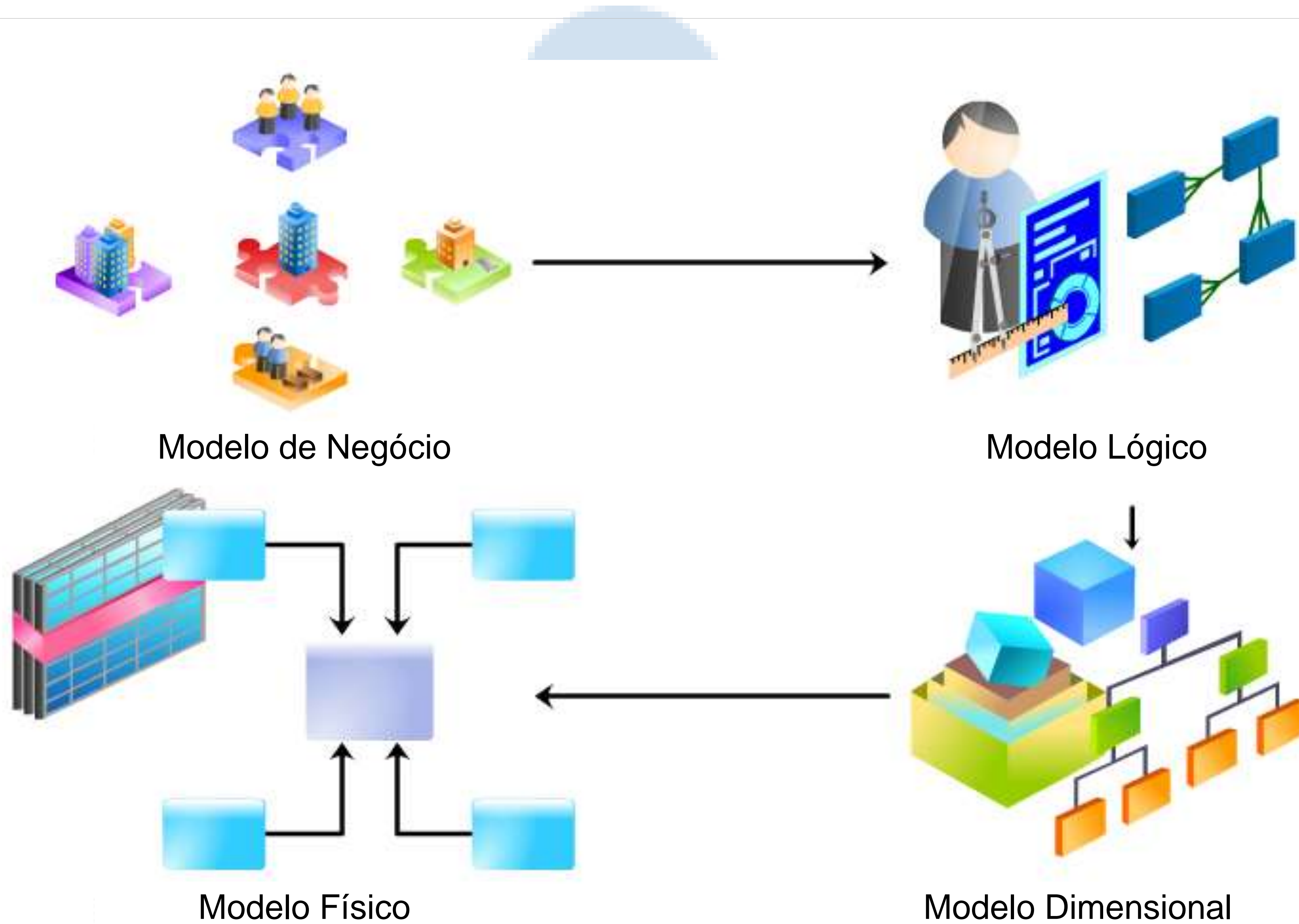
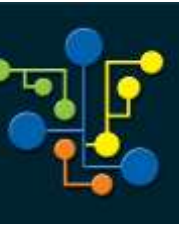


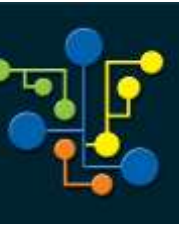
Fases de Design de um Data Warehouse





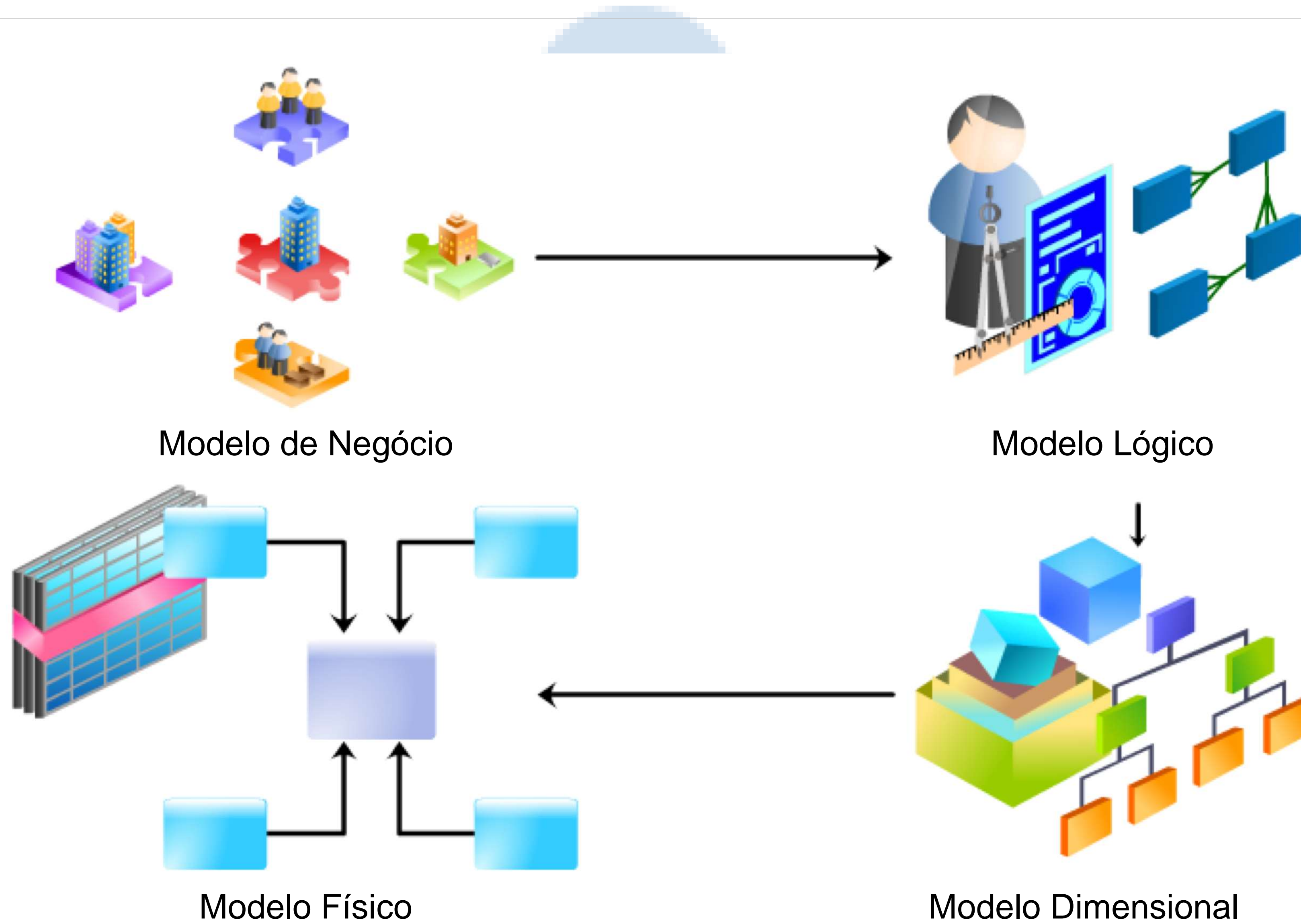






Definindo o Modelo de Negócio

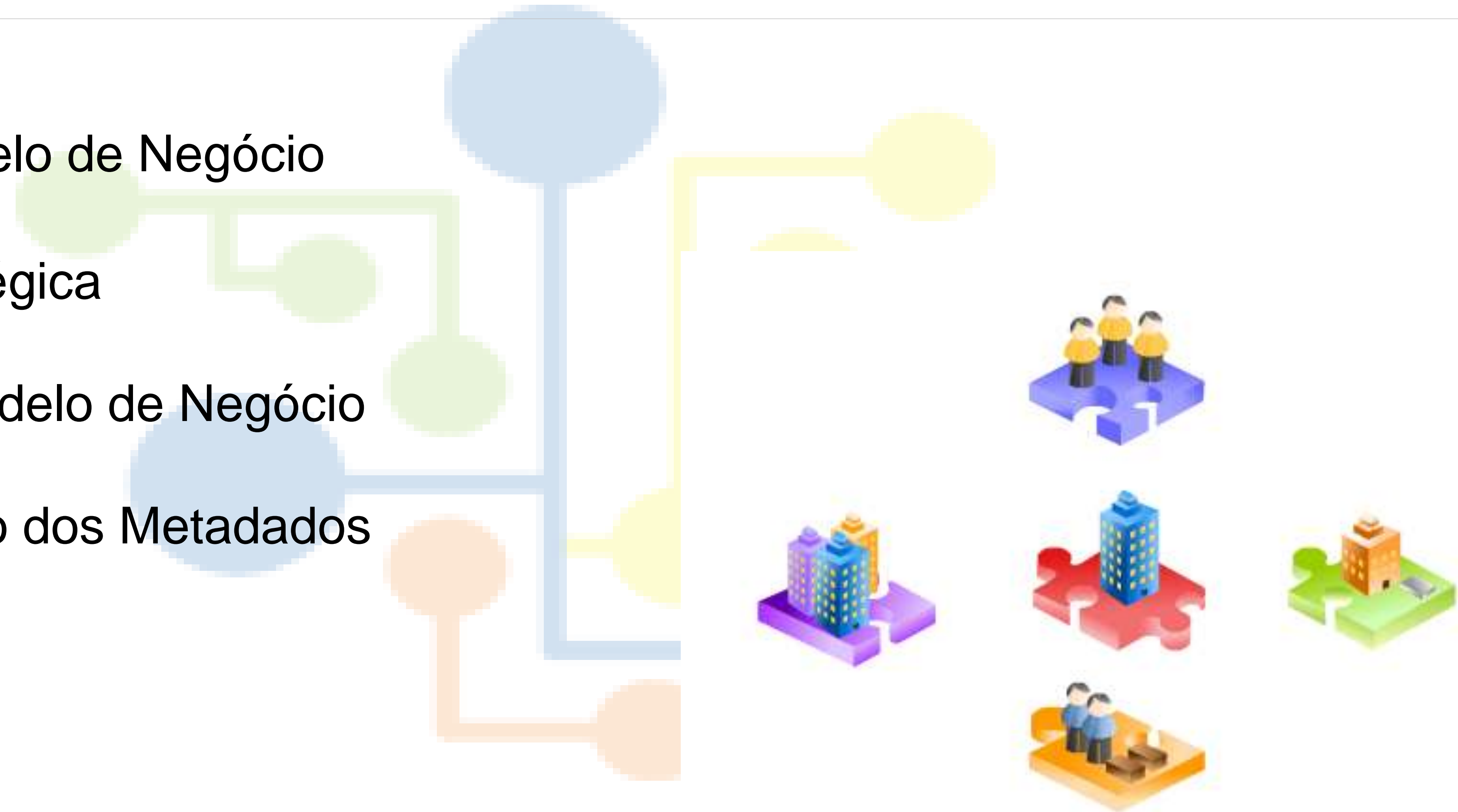






Definindo o Modelo de Negócio

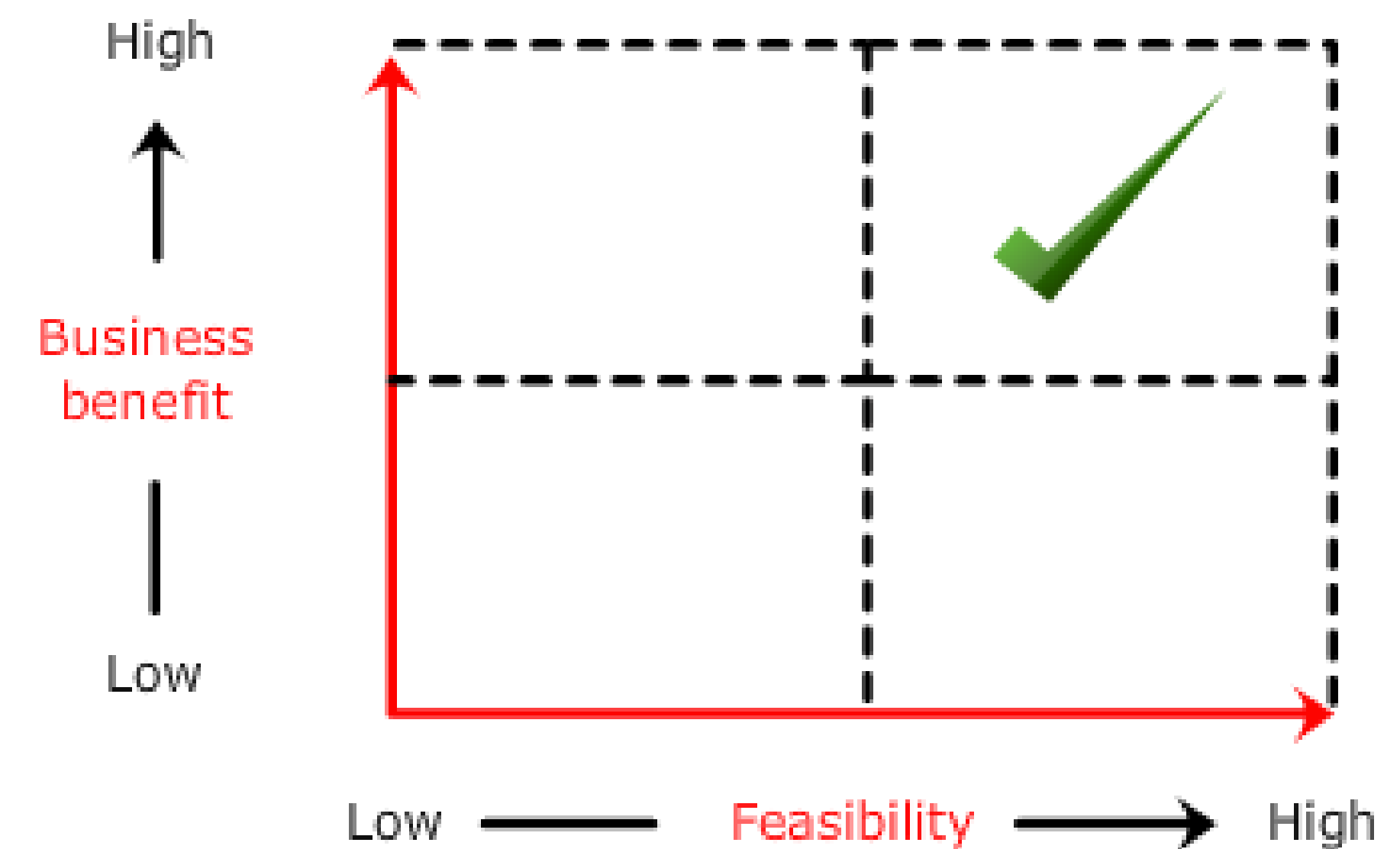
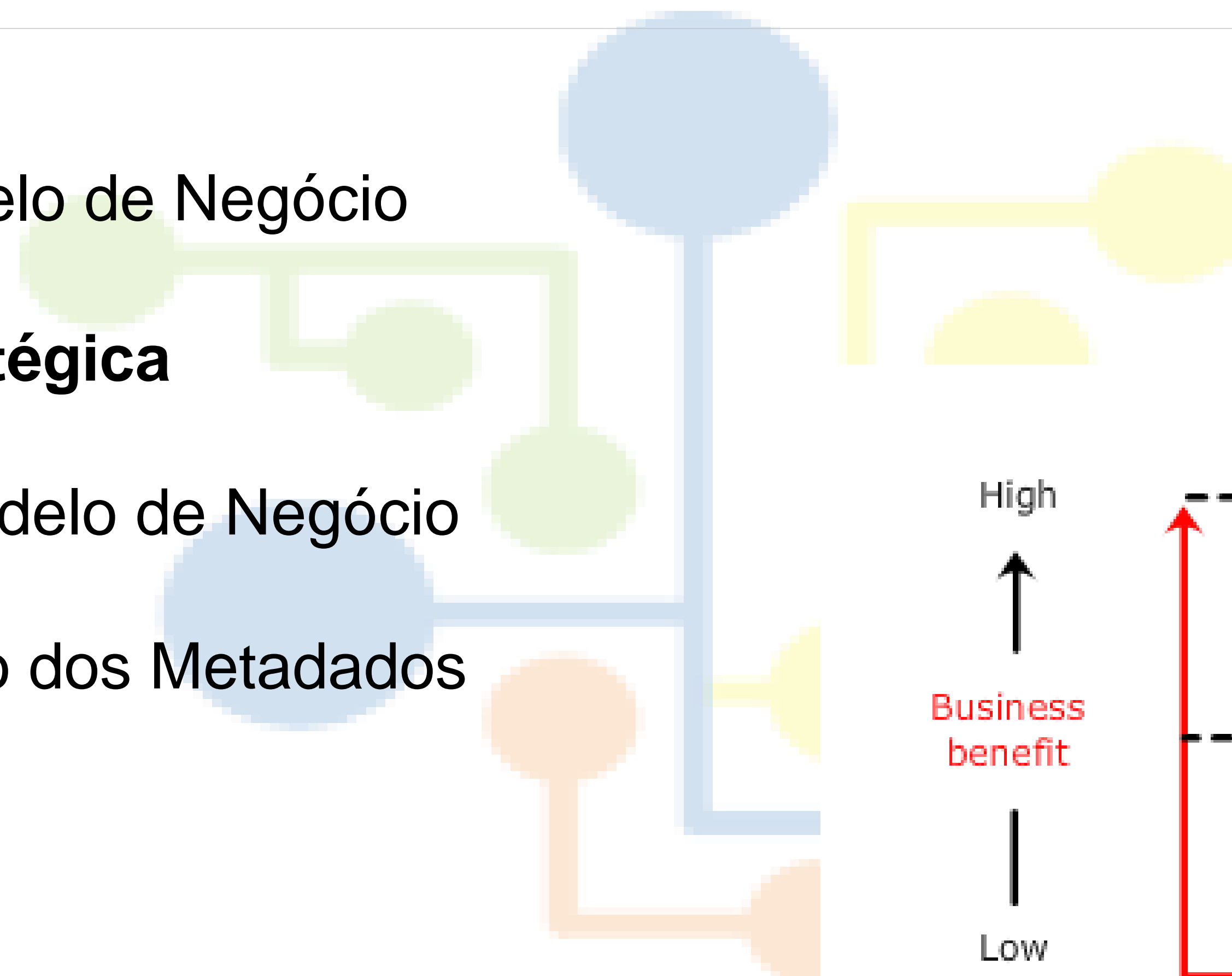
- Análise Estratégica
- Criação do Modelo de Negócio
- Documentação dos Metadados





Definindo o Modelo de Negócio

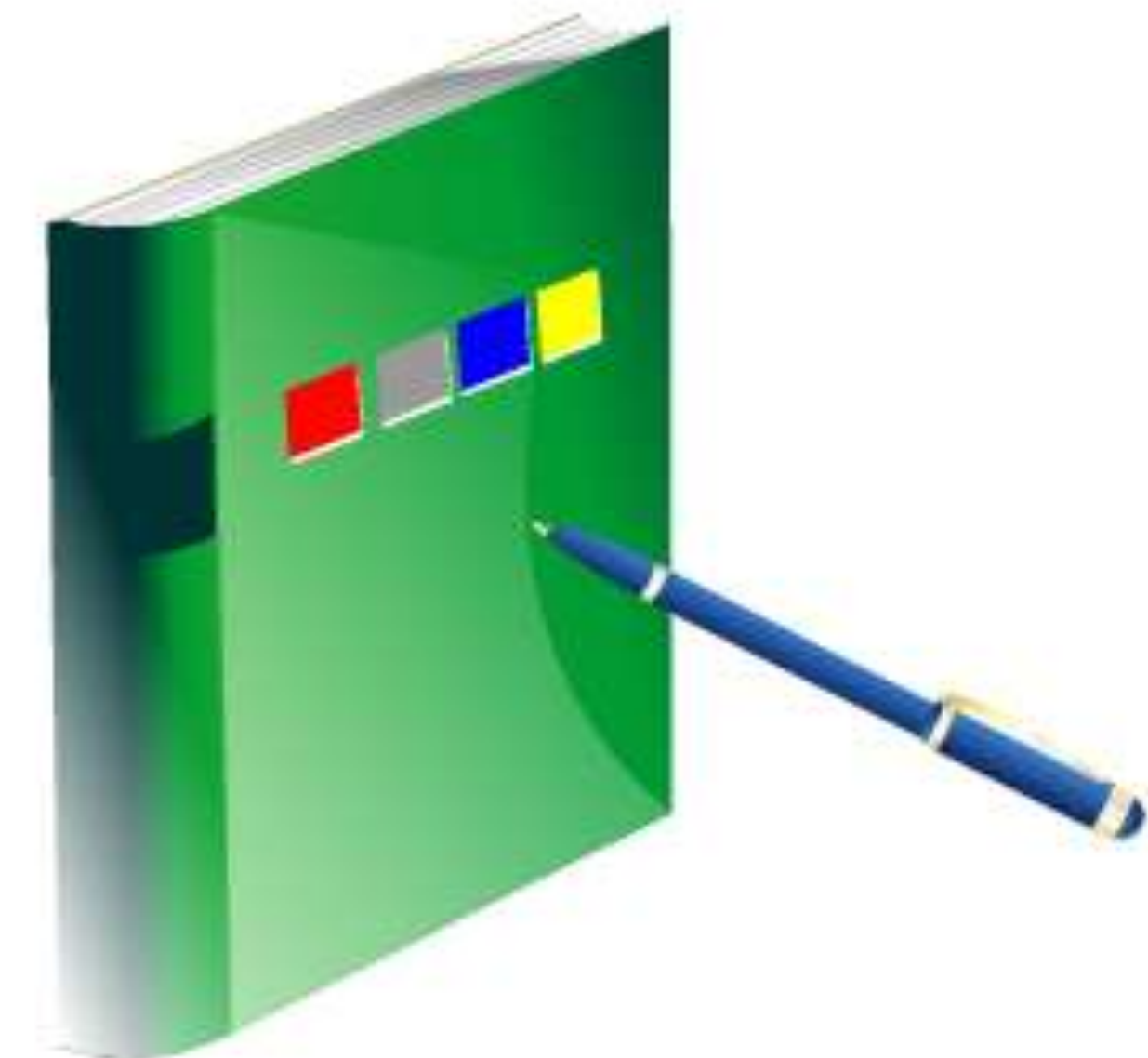
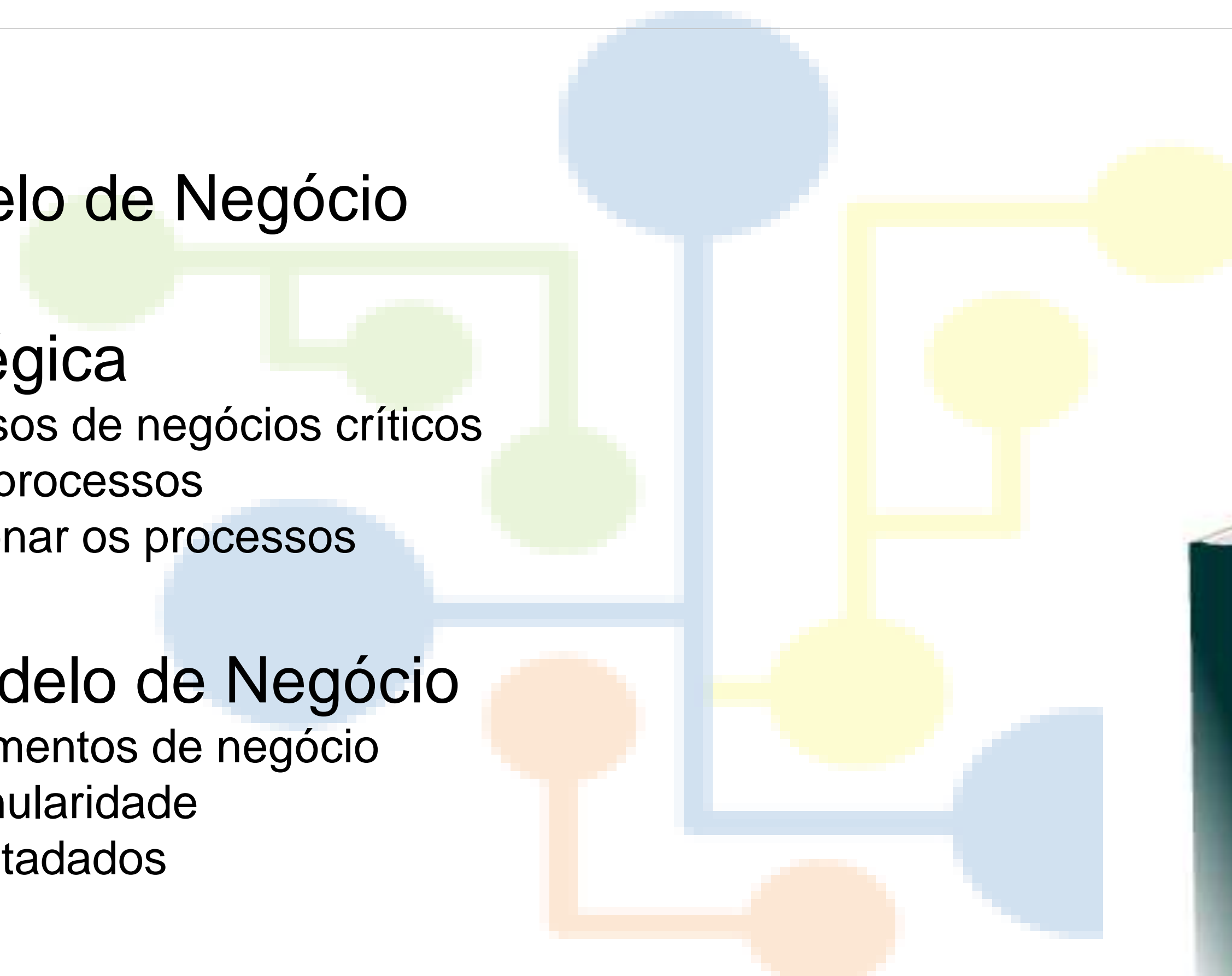
- **Análise Estratégica**
- Criação do Modelo de Negócio
- Documentação dos Metadados

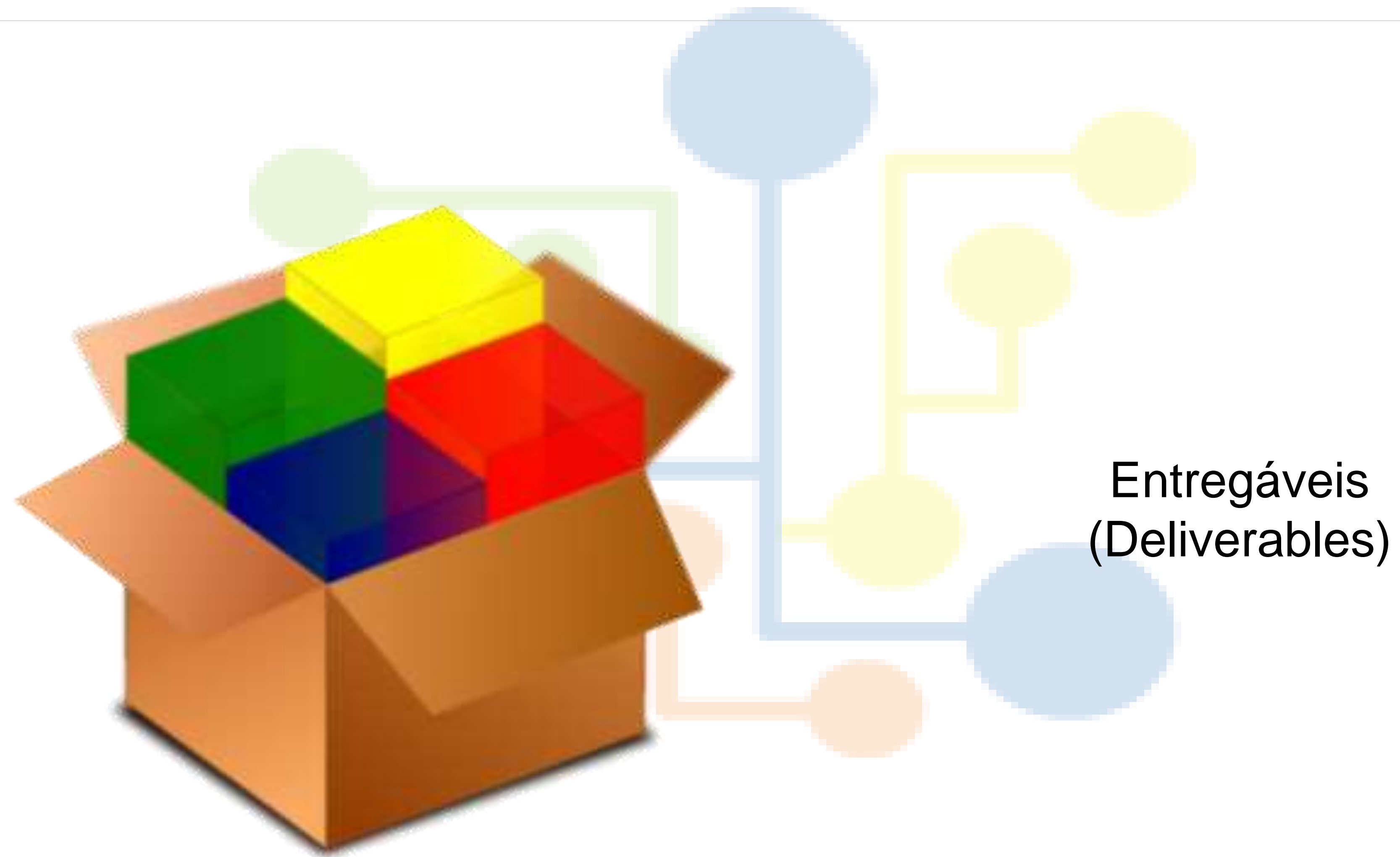




Definindo o Modelo de Negócio

- **Análise Estratégica**
 - Identificar processos de negócios críticos
 - Compreender os processos
 - Priorizar e selecionar os processos
- **Criação do Modelo de Negócio**
 - Definir os requerimentos de negócio
 - Determinar a granularidade
 - Determinar os metadados
- **Documentação dos Metadados**
 - Documentar o design do processo
 - Documentar registros de mudanças
 - Registrar melhorias ao longo do tempo







Identificar Dimensões e Medidas

Dimensões (atributos percebidos como constantes)

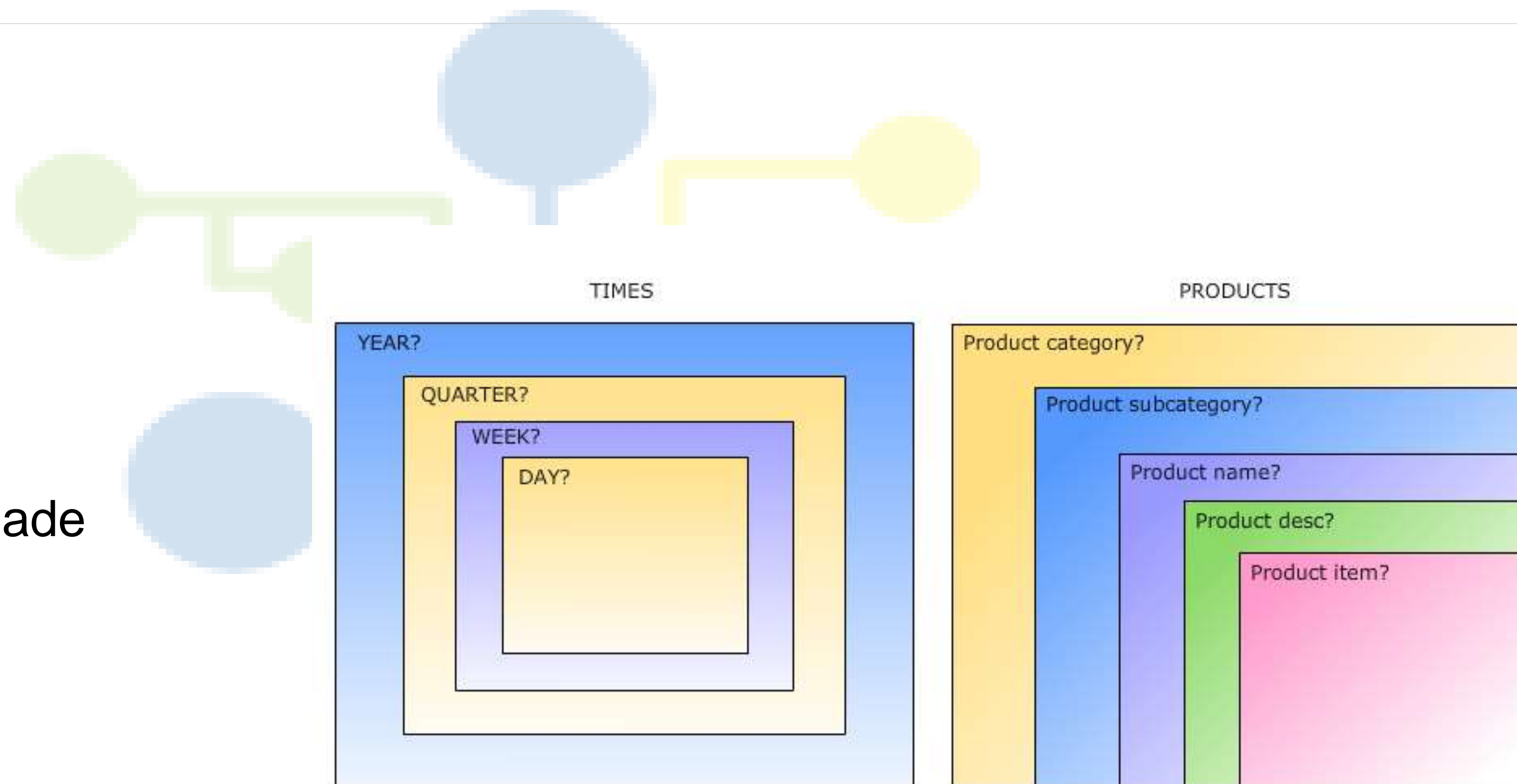
- Produtos
- Promoções
- Clientes
- Países
- Canais de Vendas
- Tempo

Medidas (atributos percebidos como variáveis)

- Vendas
- Quantidade Vendida
- Unidades Vendidas
- Custo



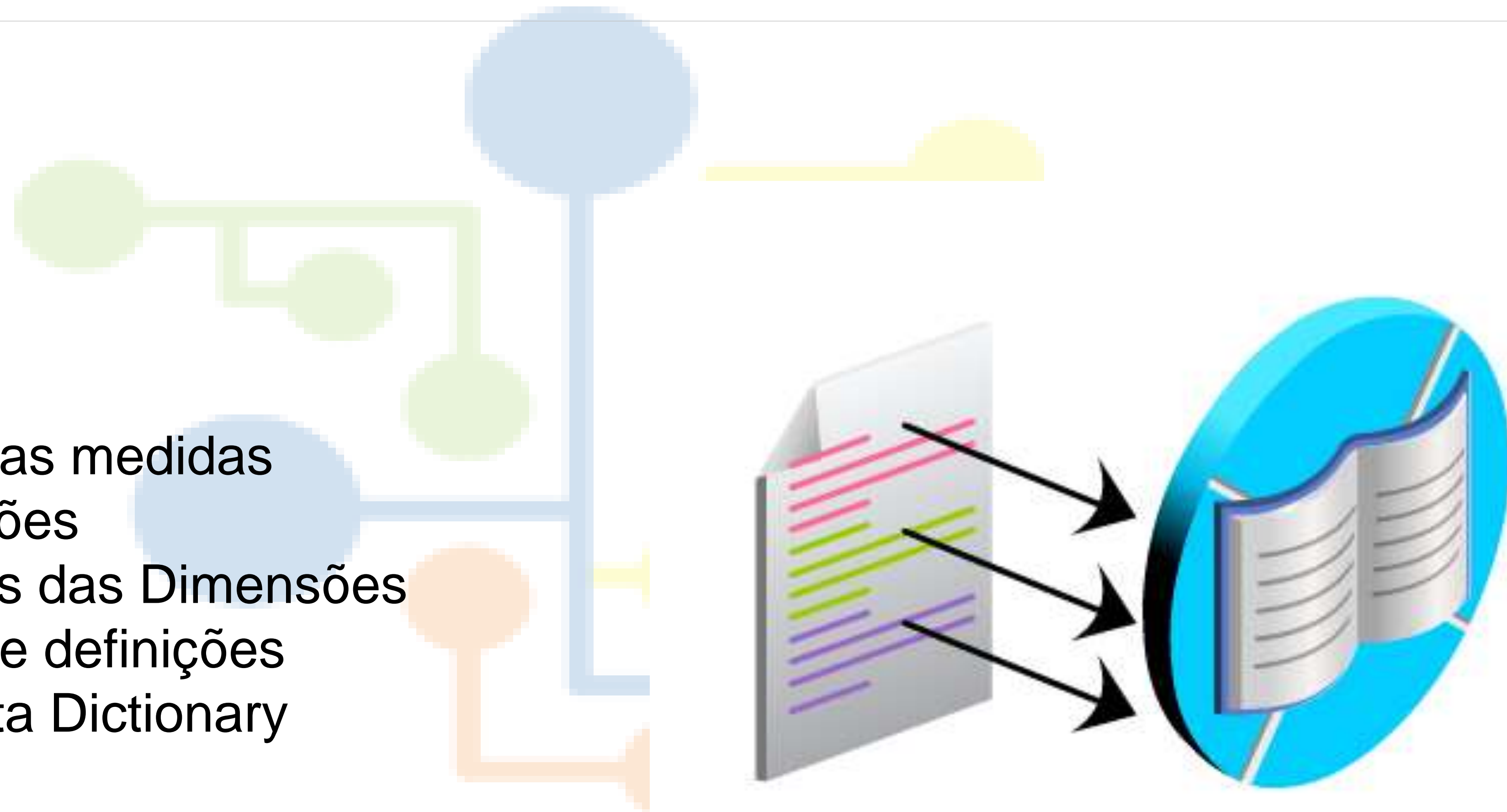
Granularidade





Metadados

- Nome das medidas
- Dimensões
- Atributos das Dimensões
- Regras e definições
- Metadata Dictionary





O resultado desta fase normalmente é uma Business Process Matrix (Matriz de Processos de Negócio)



| Business Dimensions | Business Processes | | |
|---------------------|--------------------|---------|-----------|
| | Sales | Returns | Inventory |
| Customers | ✓ | ✓ | |
| Times (Date) | ✓ | ✓ | ✓ |
| Products | ✓ | ✓ | ✓ |
| Channels | ✓ | | |
| Promotions | ✓ | ✓ | |

Sample of business process matrix



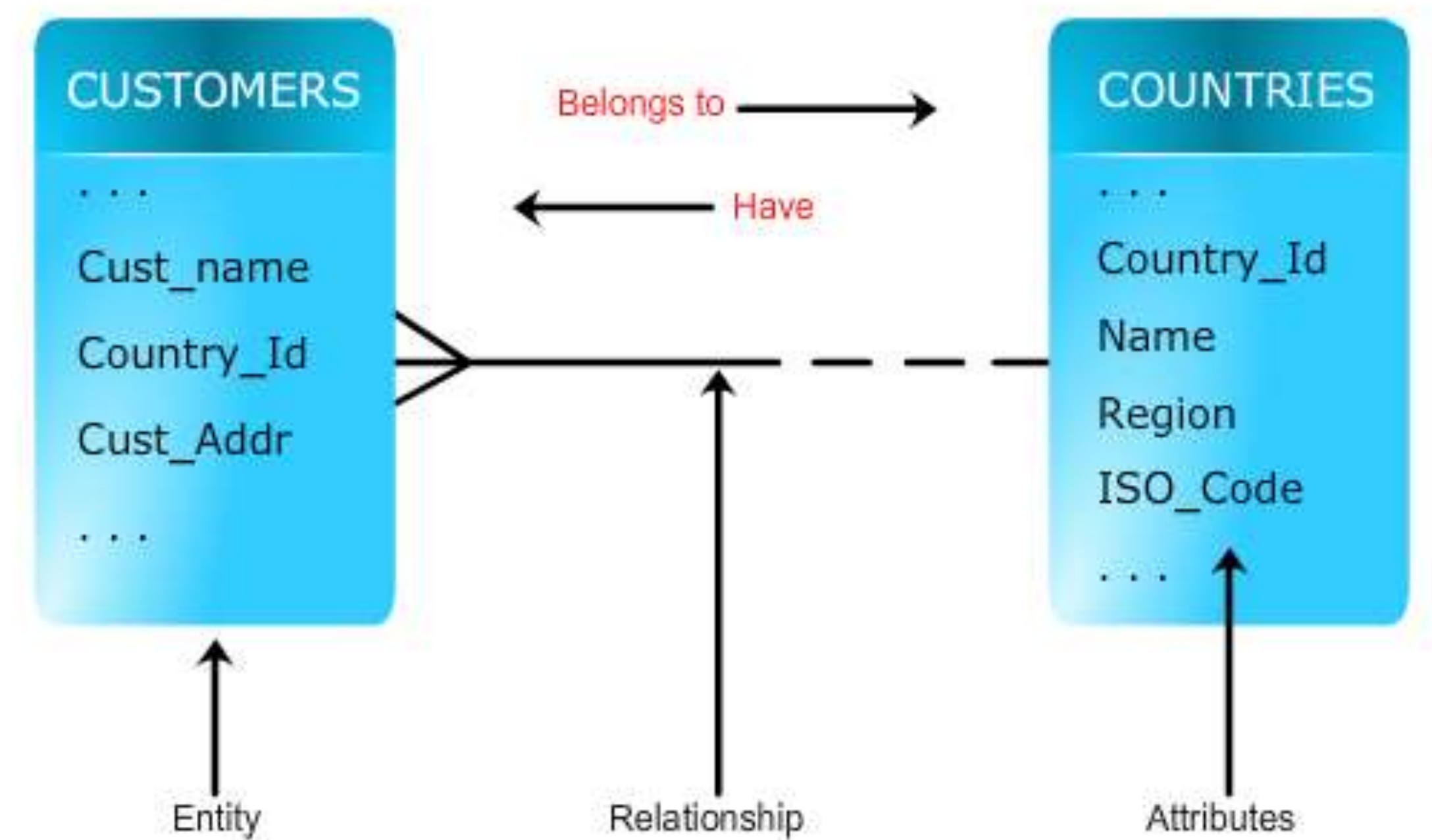
Definindo o Modelo Lógico

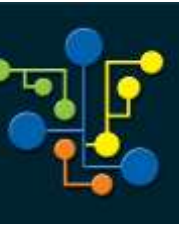




Definindo o Modelo Lógico

- Modelo Entidade-Relacionamento
- Diagrama Entidade-Relacionamento





Definindo o Modelo Dimensional

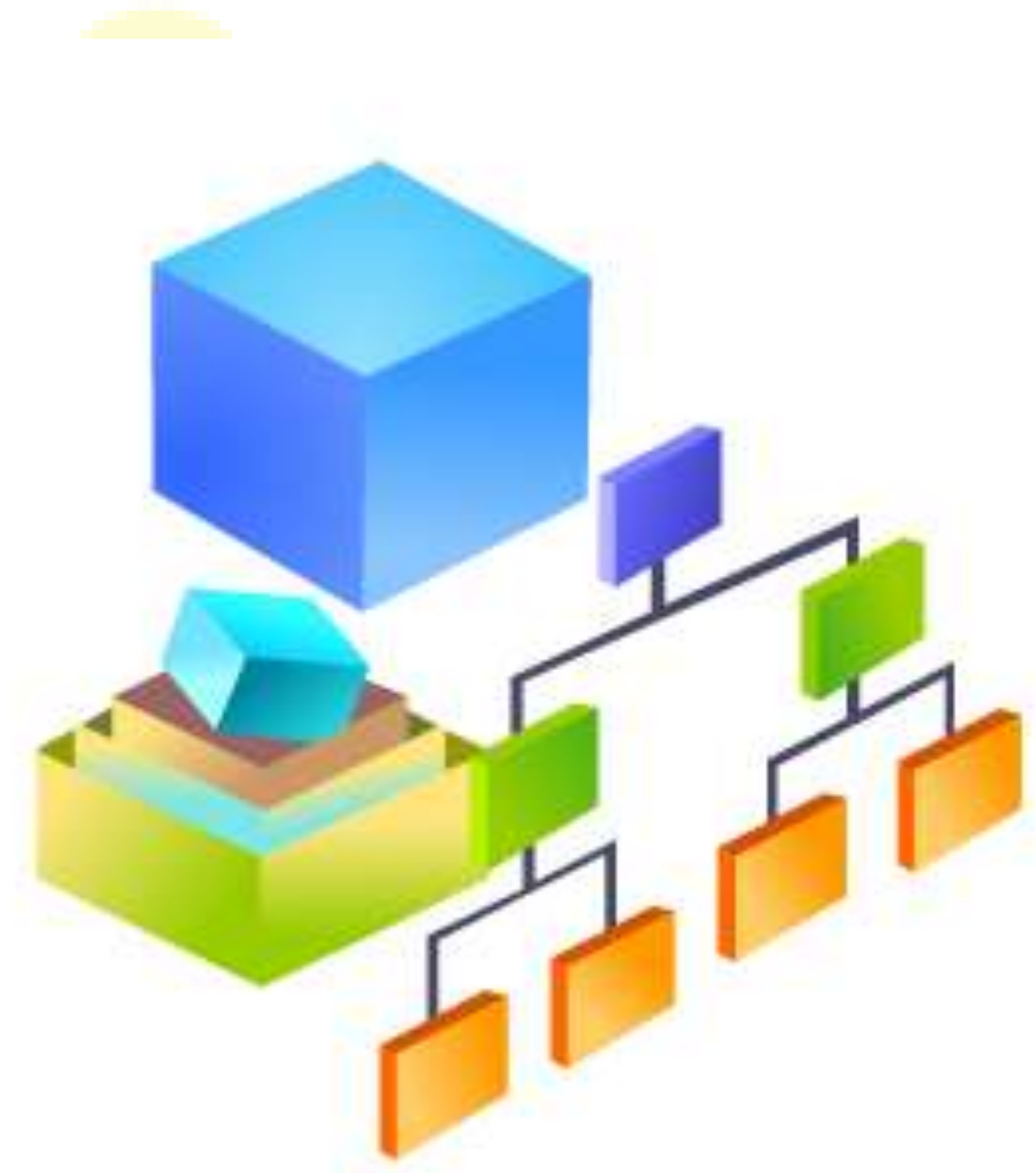


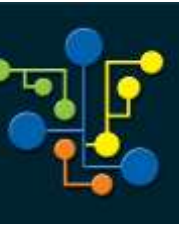


Definindo o Modelo Dimensional

O Modelo Lógico (criado na fase anterior) é fundamental na fase de Modelagem Dimensional, pois ajuda a:

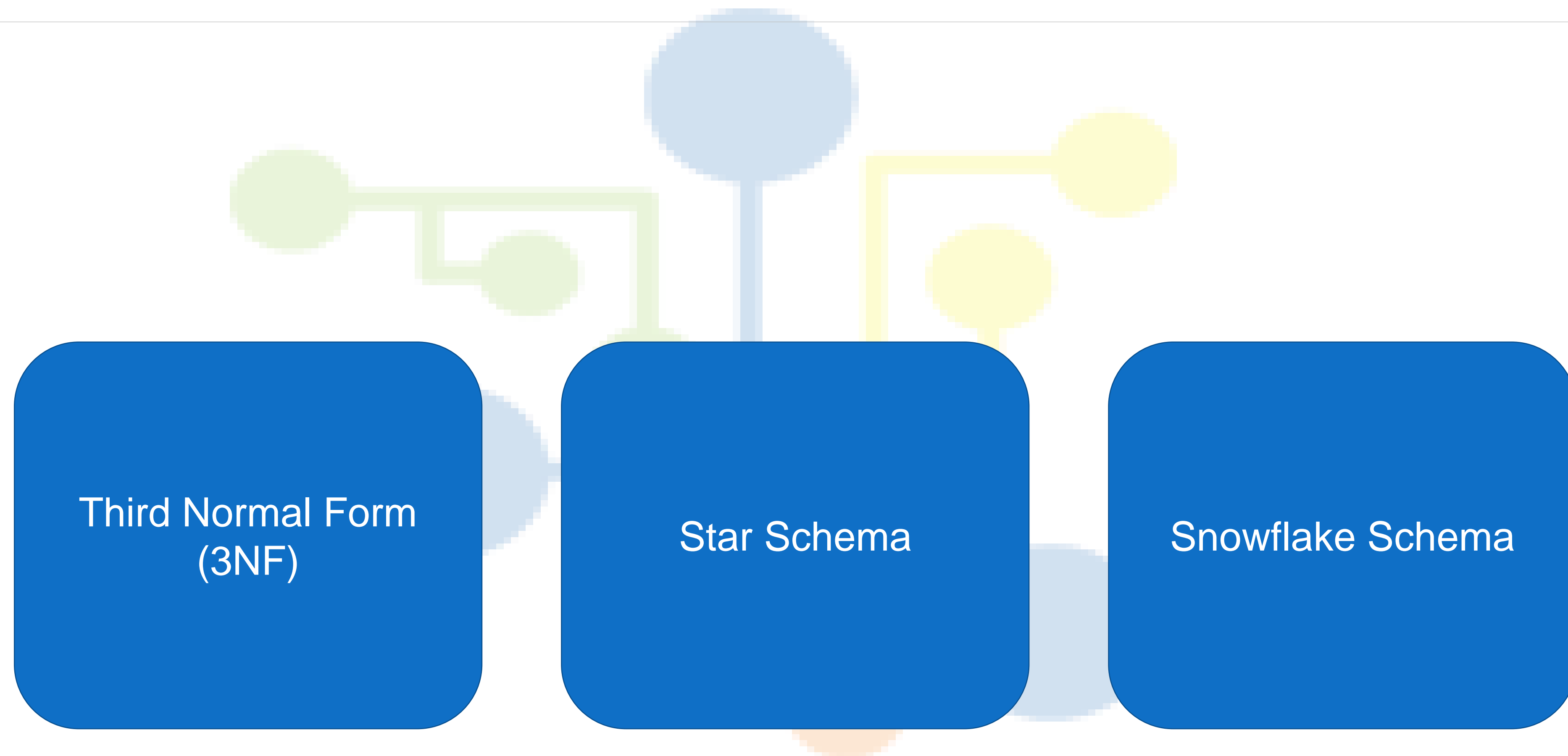
- Identificar a tabela FATO
- Identificar tabelas DIMENSÃO
- Estabelecer links entre Fatos e Dimensões
- Modelar Dimensão Tempo





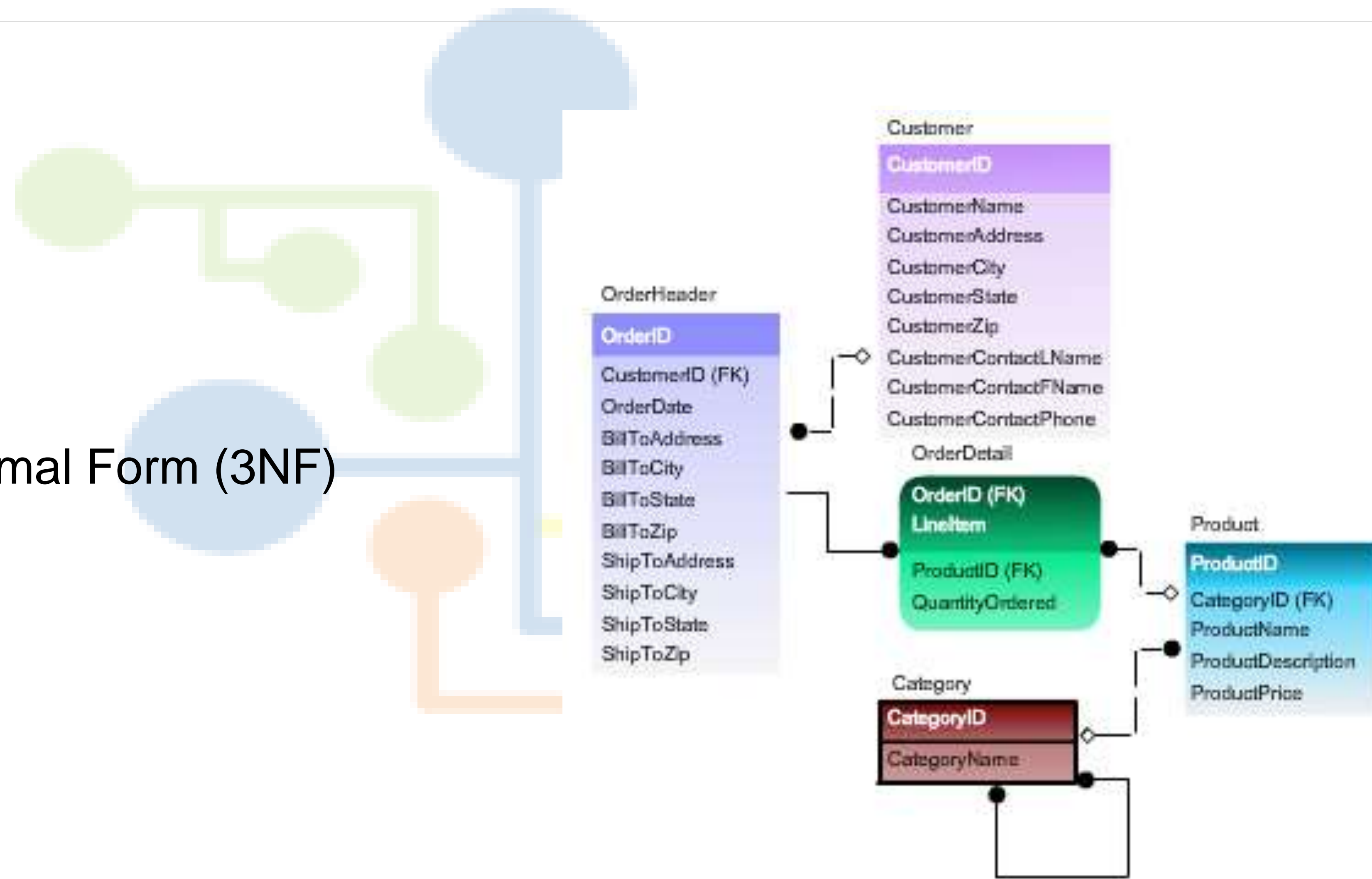
Data Warehouse Schemas





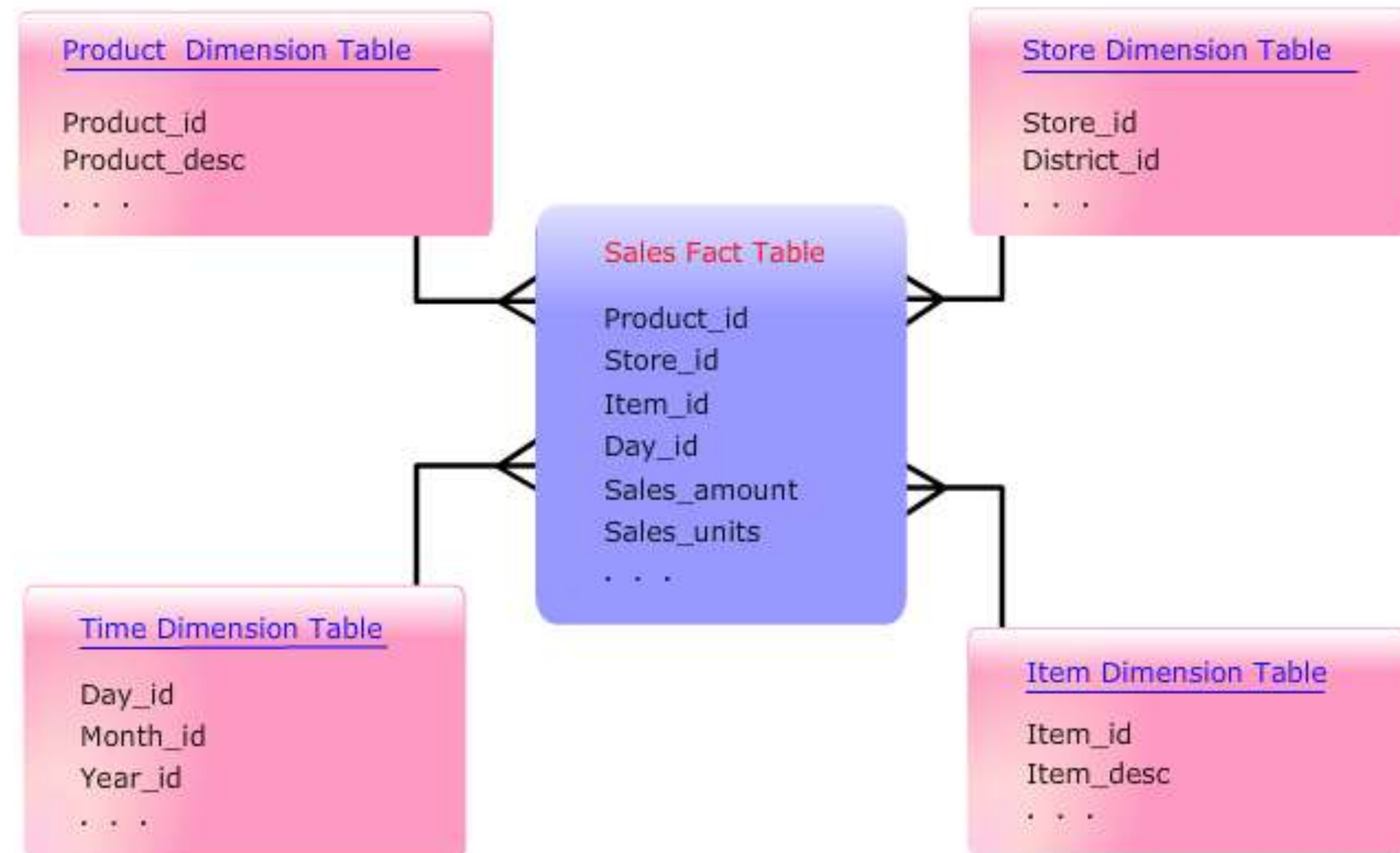


Third Normal Form (3NF)



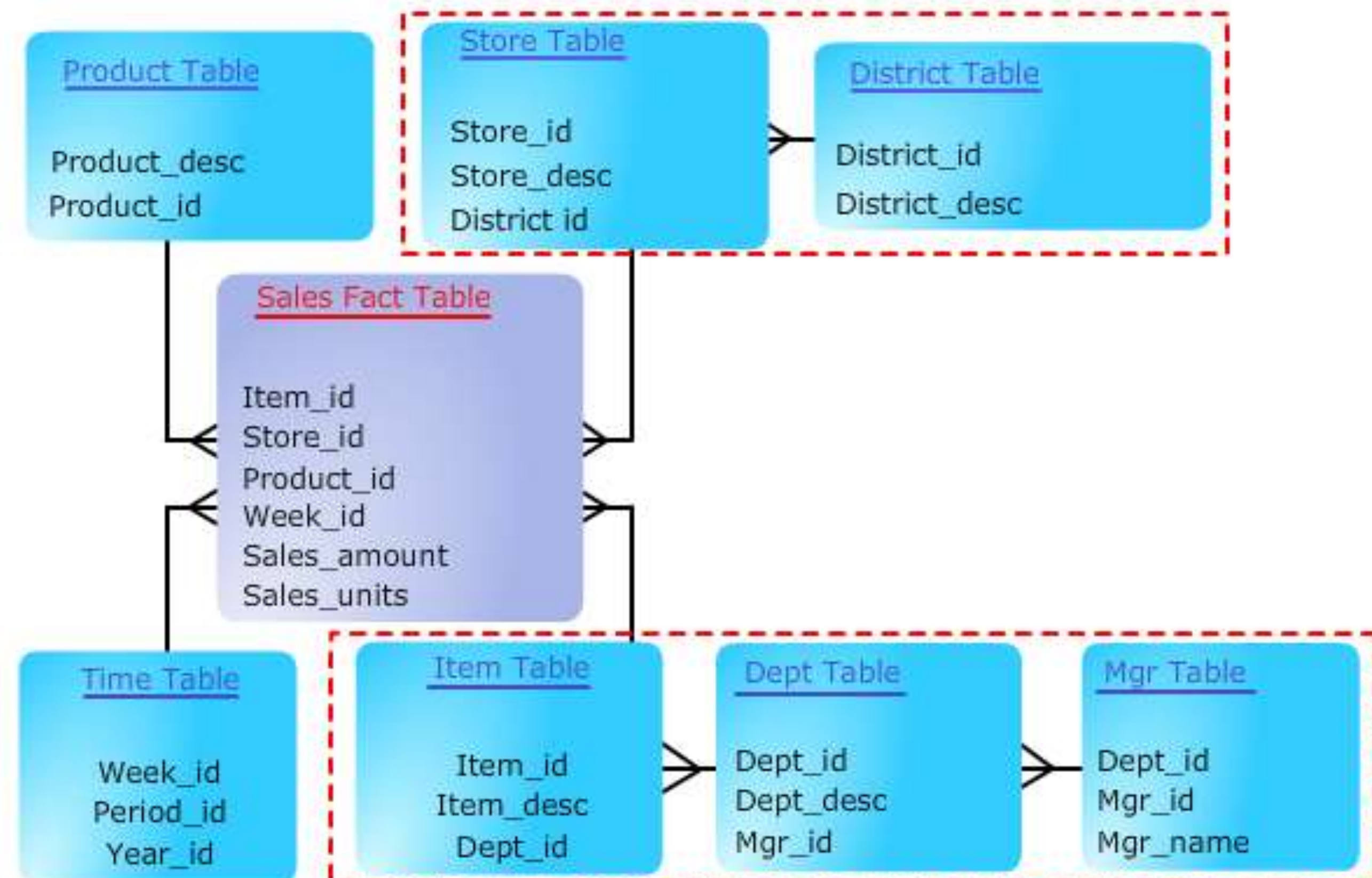


Star Schema





Snowflake Schema





Tabelas Fato e Dimensão





Tabela Fato

- Contém métricas numéricas do negócio
- Tabela com maior volume de dados
- Cresce rapidamente
- Pode conter dados base, derivados e sumarizados
- São relacionadas às tabelas DIMENSÃO com chaves estrangeiras que referenciam as chaves primárias nas tabelas DIMENSÃO

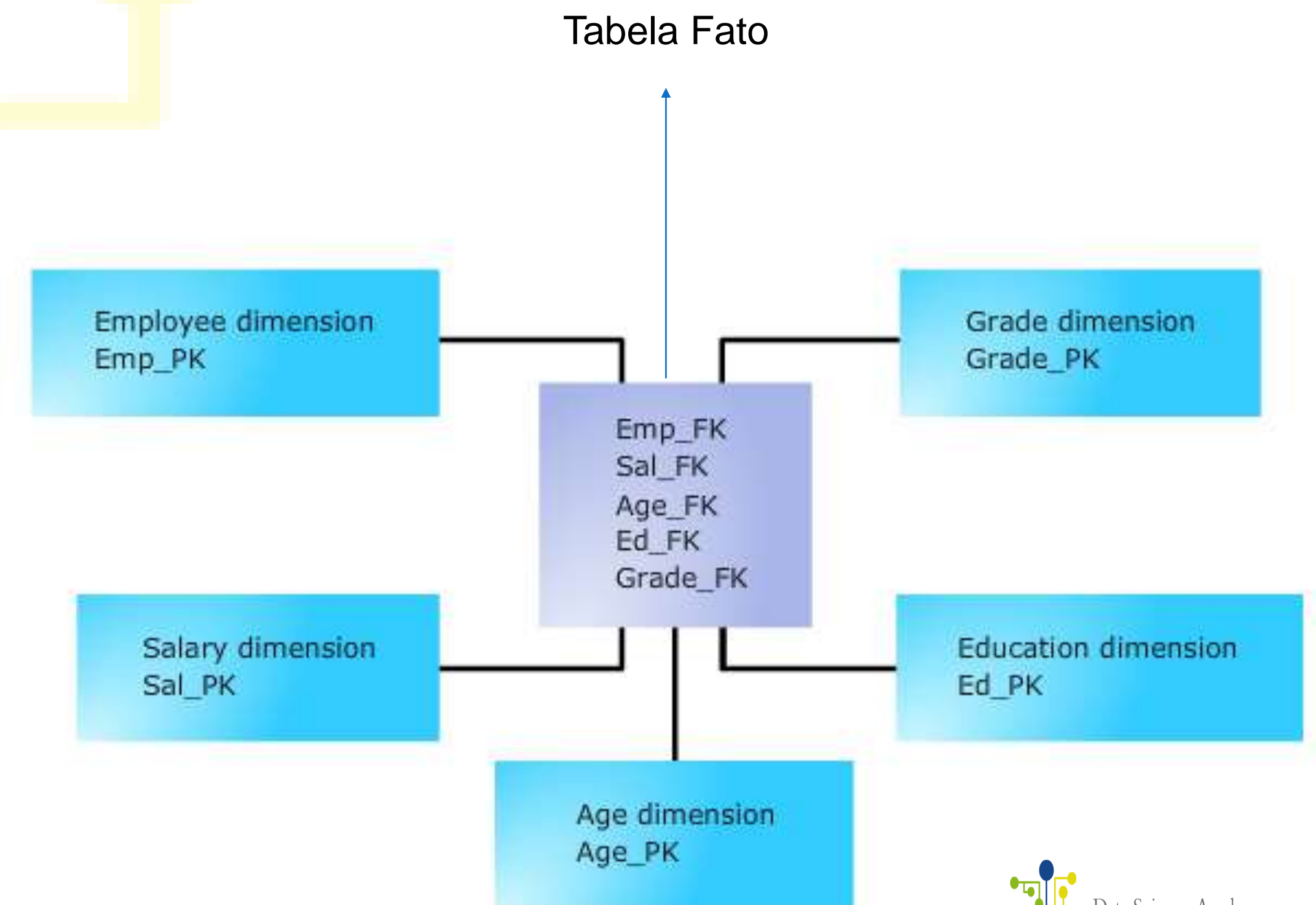
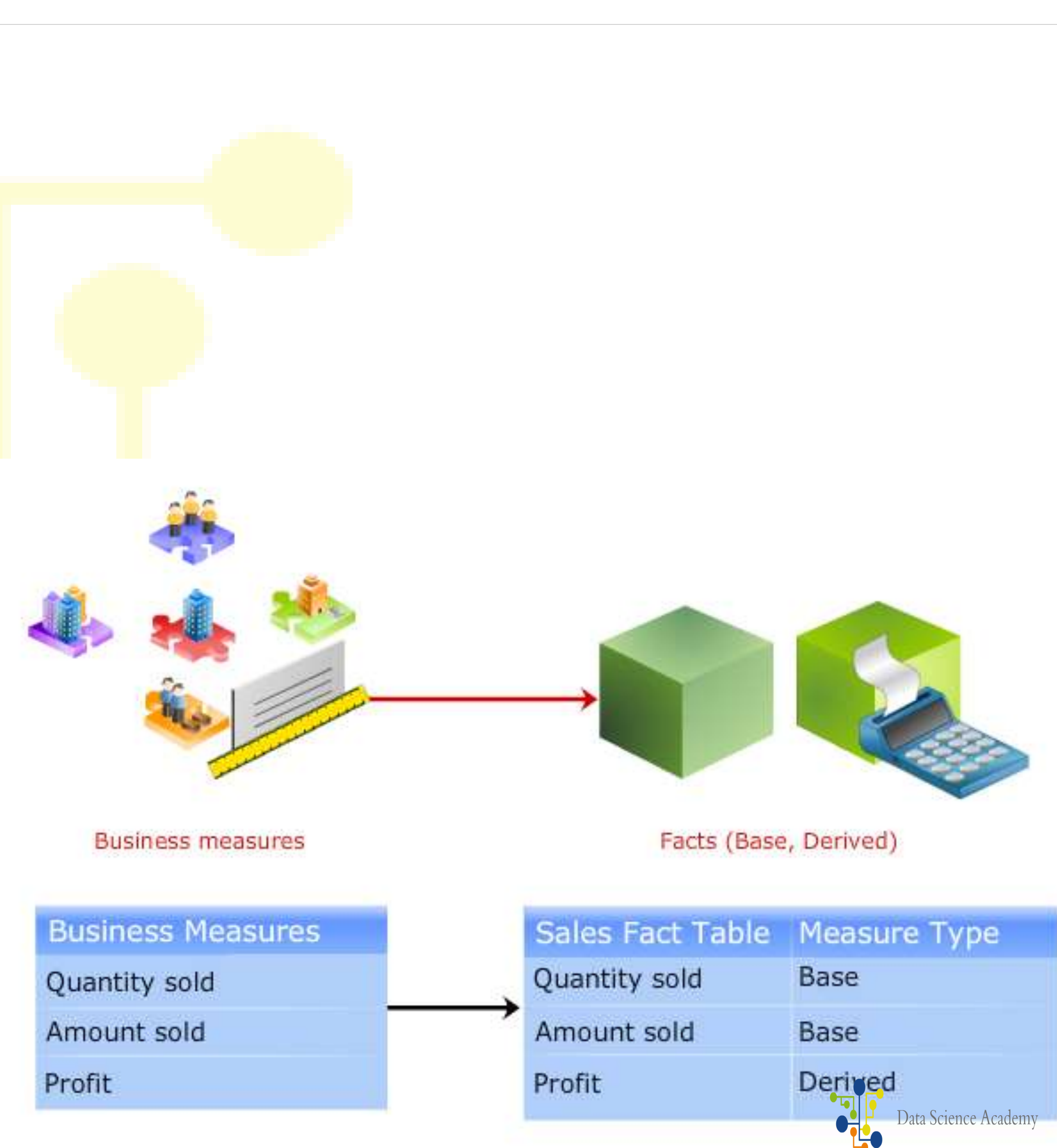




Tabela Fato

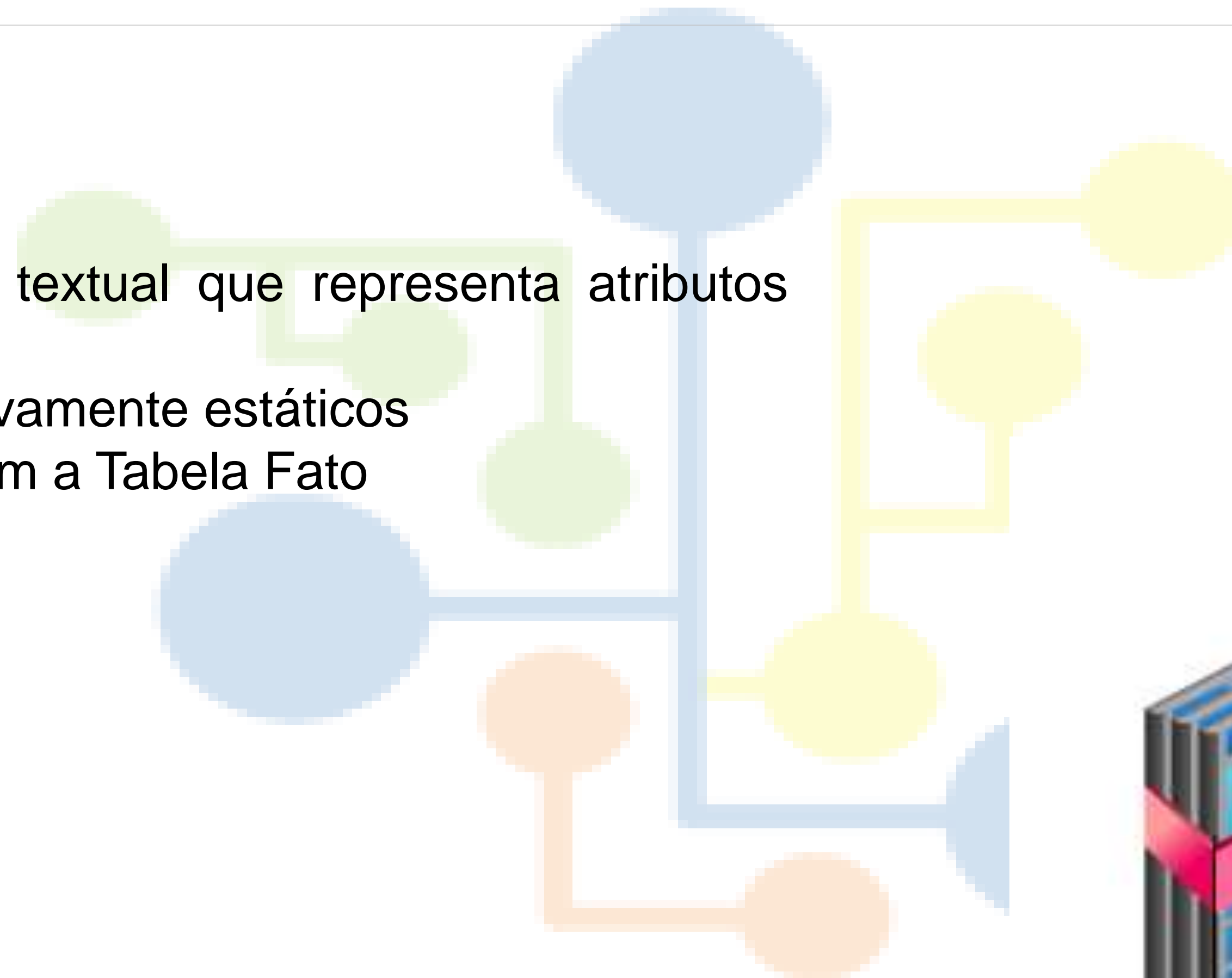
- Contém métricas numéricas do negócio
- Tabela com maior volume de dados
- Cresce rapidamente
- Pode conter dados base, derivados e sumarizados
- São relacionadas às tabelas DIMENSÃO com chaves estrangeiras que referenciam as chaves primárias nas tabelas DIMENSÃO





Tabelas DIMENSÃO

- Contém informação textual que representa atributos de negócio
- Contém dados relativamente estáticos
- São relacionadas com a Tabela Fato





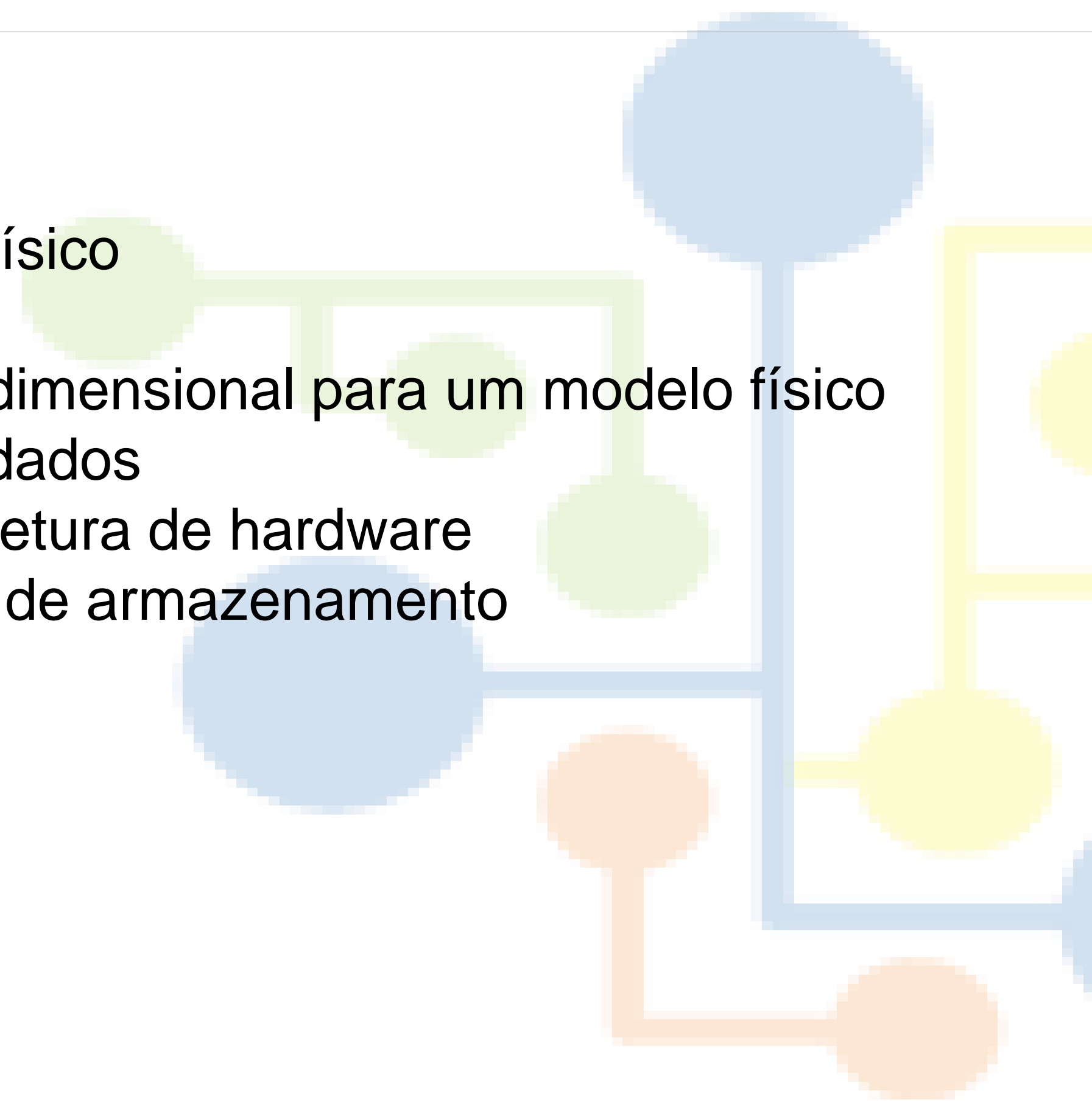
Definindo o Modelo Físico





Definindo o Modelo Físico

- Traduz do design dimensional para um modelo físico
- Atualiza dos metadados
- Determina a arquitetura de hardware
- Define estratégias de armazenamento
- Sizing
- Particionamento
- Indexação
- Segurança



| | |
|-----------------|---|
| # *Product_PK | n |
| # *Channel_PK | n |
| # *Promotion_PK | n |

| | |
|---------------------|--------|
| *PRODUCT_ID | v(11) |
| *PROD_DESC | v(125) |
| *PROD_NAME | v(35) |
| *PROD_CATEGORY_ID | v(20) |
| *PROD_CATEGORY_DESC | v(50) |
| *PROD_SUBCATEGORY | v(25) |
| *SUPPLIER_ID | v(20) |
| *PROD_STATUS | v(10) |
| *PROD_LIST_PRICE | n |
| *PROD_MIN_PRICE | v(20) |
| *PROD_PACK_SIZE | v(20) |
| *PROD_WEIGHT_CLASS | v(10) |
| *PROMOTION_CODE | v(10) |
| *WHSE_LOCATION | v(10) |
| *PROD_EFF_FROM | date |
| *PROD_EFF_TO | date |





Muito Obrigado!

É um prazer ter você aqui.
Tenha uma excelente jornada de aprendizagem.



Oportunidade



Disponibilidade



Conhecimento