



Data Science
Academy

Data Science Academy eng.davidborges@gmail.com 59532d8f5e4cdead748b456a

Data Science Academy



Design e Implementação de Data Warehouses



Data Science
Academy

Data Science Academy eng.davidborges@gmail.com 59532d8f5e4cdead748b456a

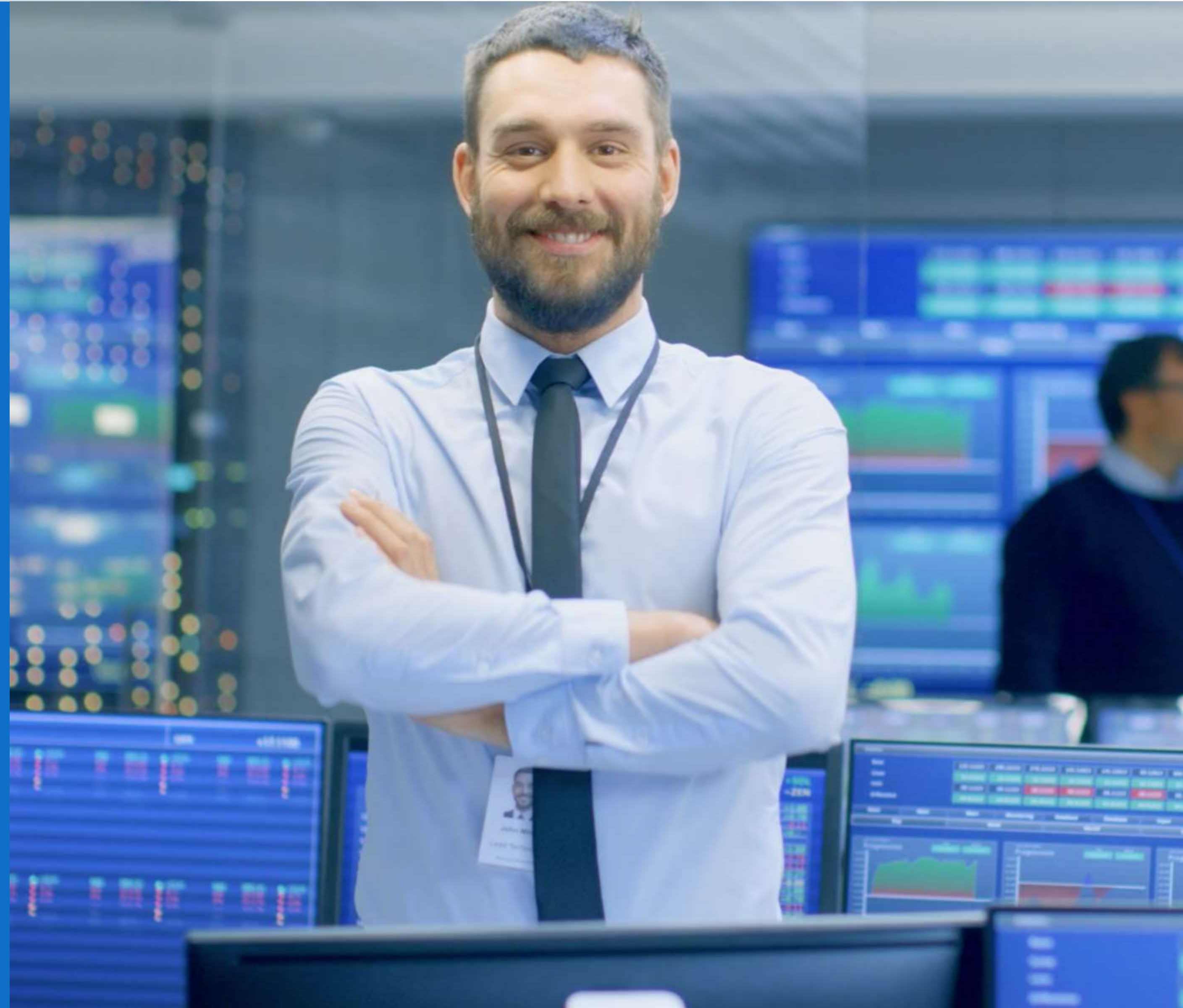
www.datascienceacademy.com.br

Design e Implementação de Data Warehouses



Data Science Academy

Projeto Implementando Data Warehouse Parte 2





Projeto





Projeto

Parte 1
Especificação

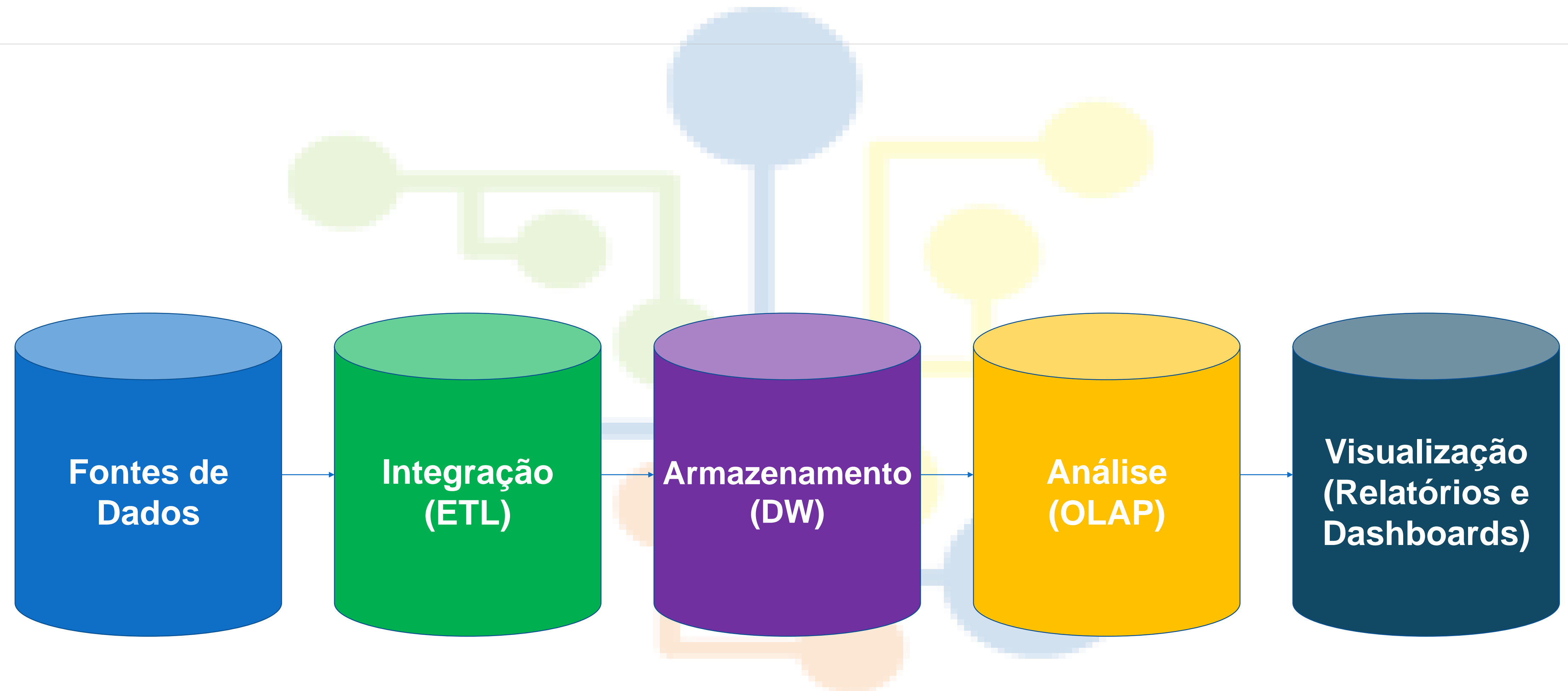
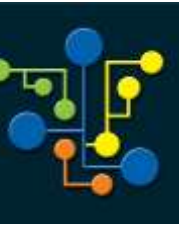
Parte 2
Implementação

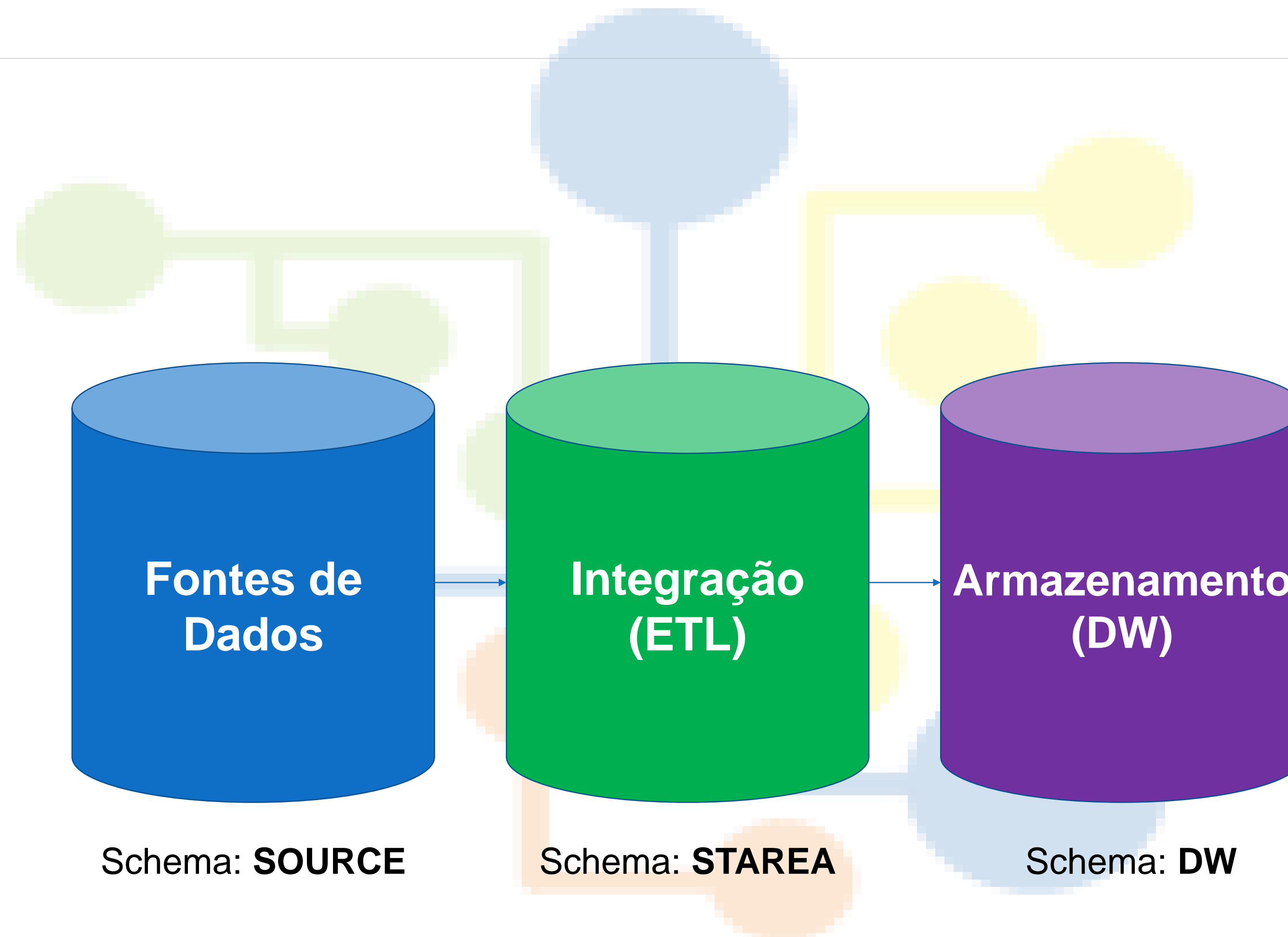
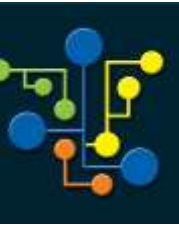




Arquitetura da Solução









Data Science Academy

Refresh do DW



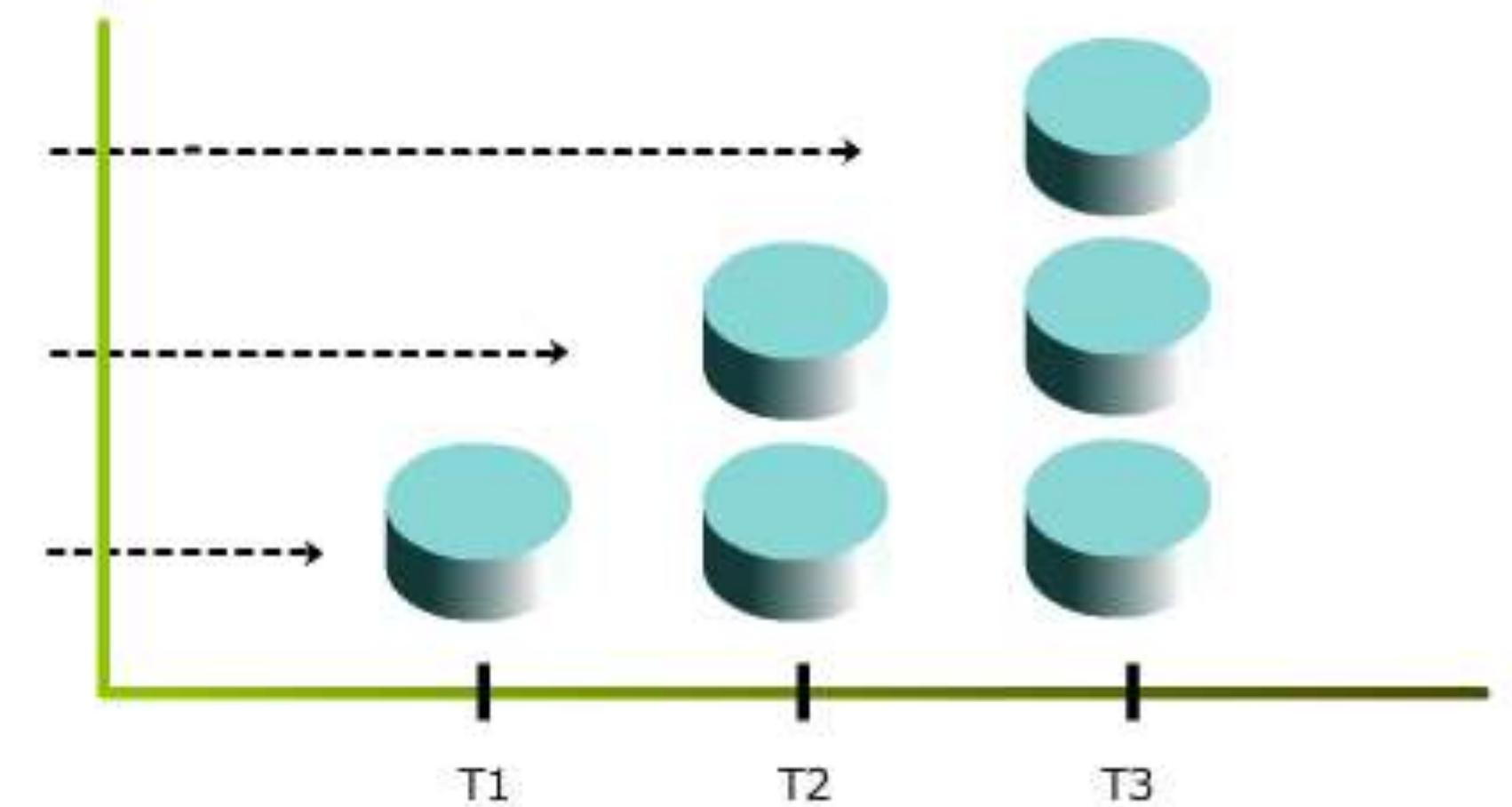


É preciso definir uma estratégia de atualização junto com a área de negócio.

- Definir a janela de atualização
- Identificar o volume de dados
- Definir o ciclo de atualização
- Planejar a Staging Area
- Determinar como detectar mudanças nos dados

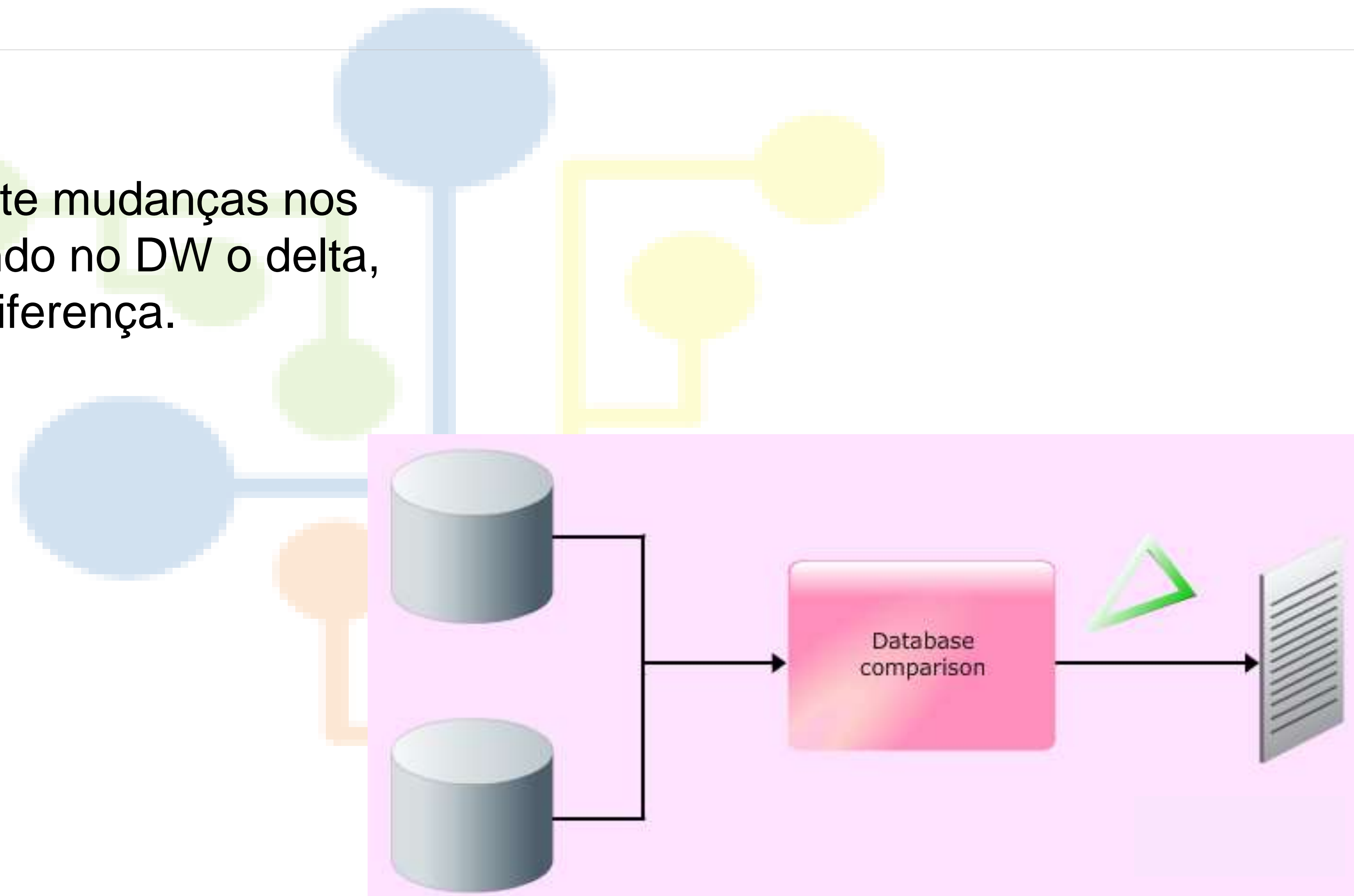


Operational
databases



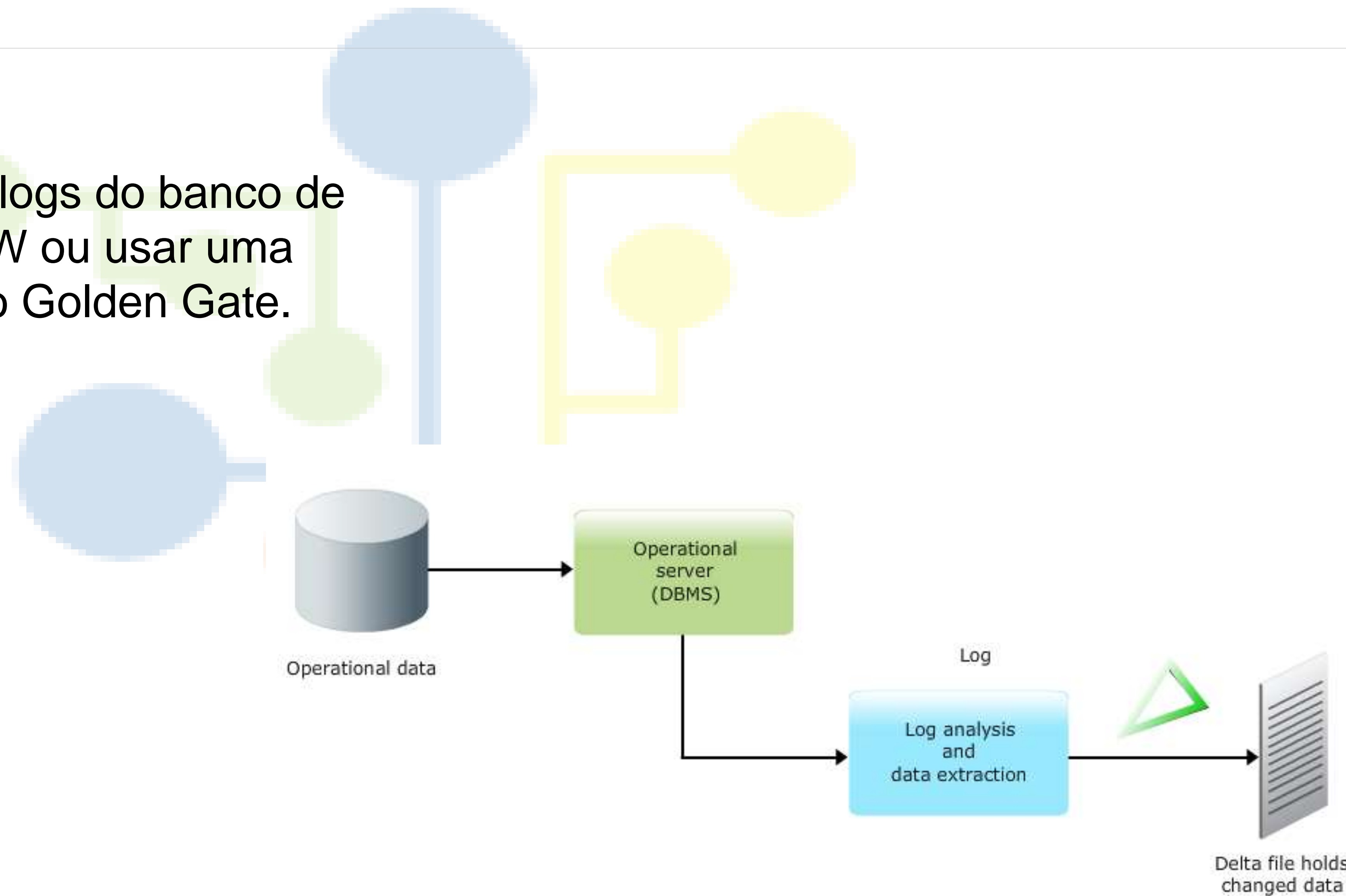


Comparar diariamente mudanças nos sistema fonte, aplicando no DW o delta, ou seja, a diferença.



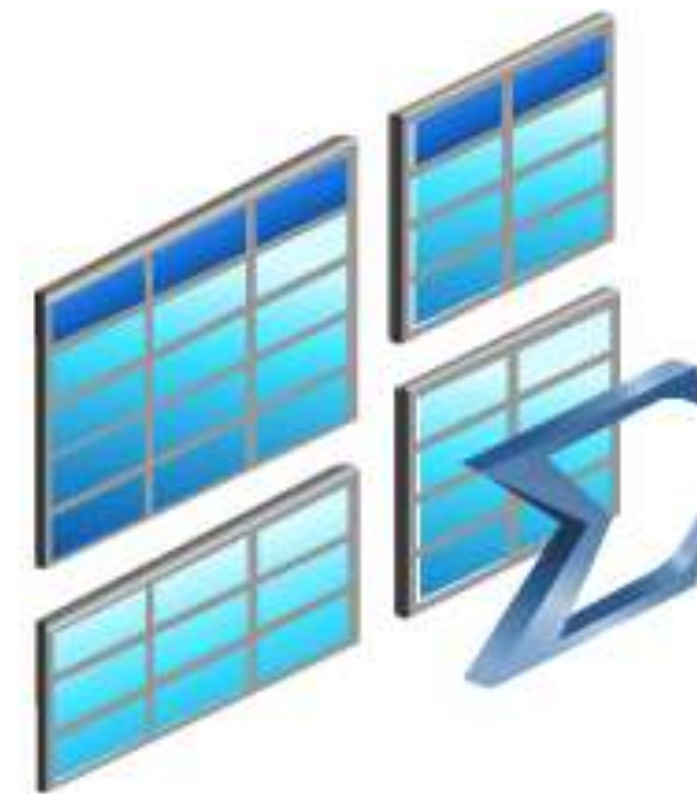
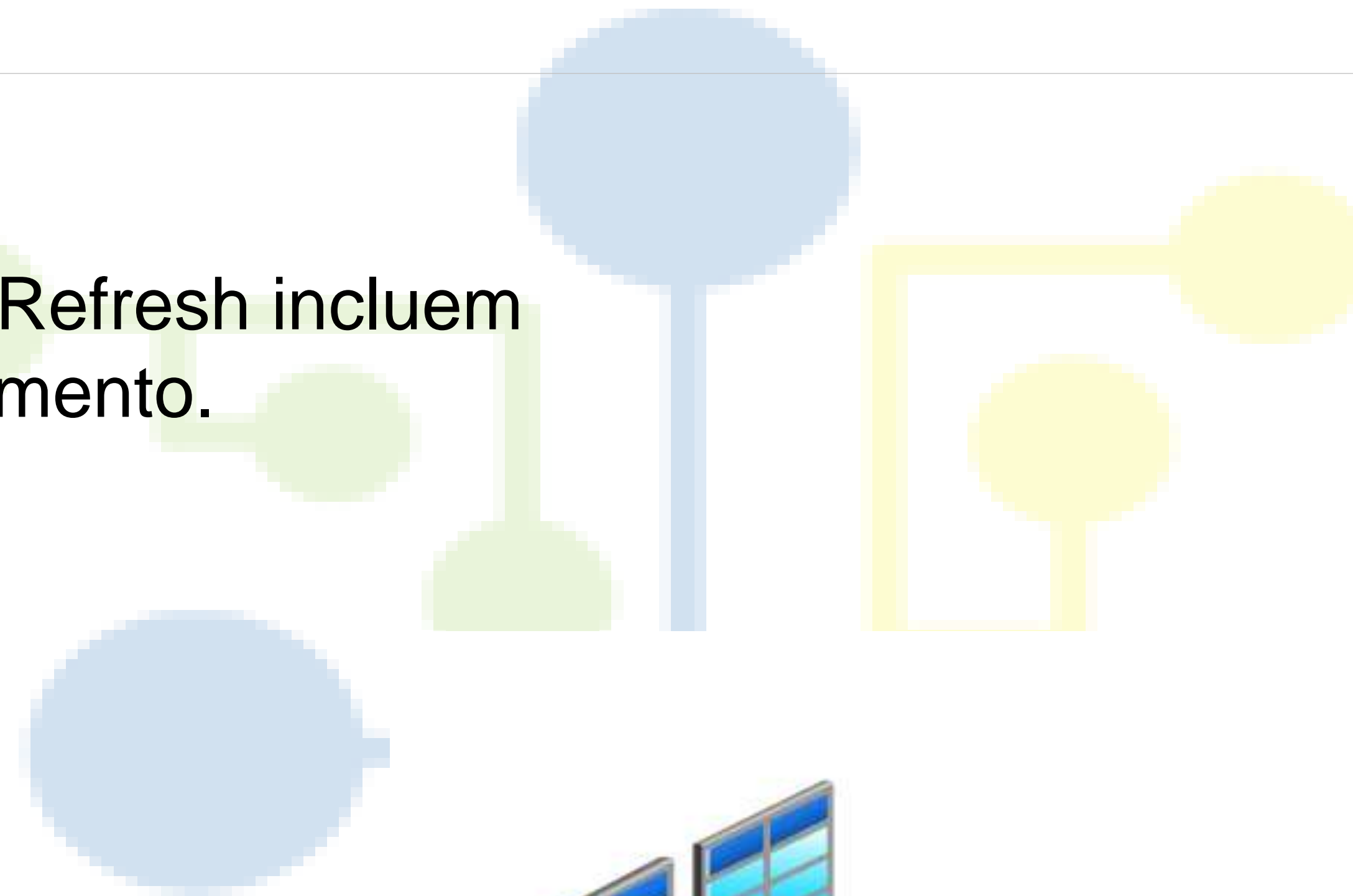


Podemos aplicar os logs do banco de dados fonte no DW ou usar uma tecnologia como o Golden Gate.





Outras técnicas de Refresh incluem
particionamento.



Use partitioning
for refresh

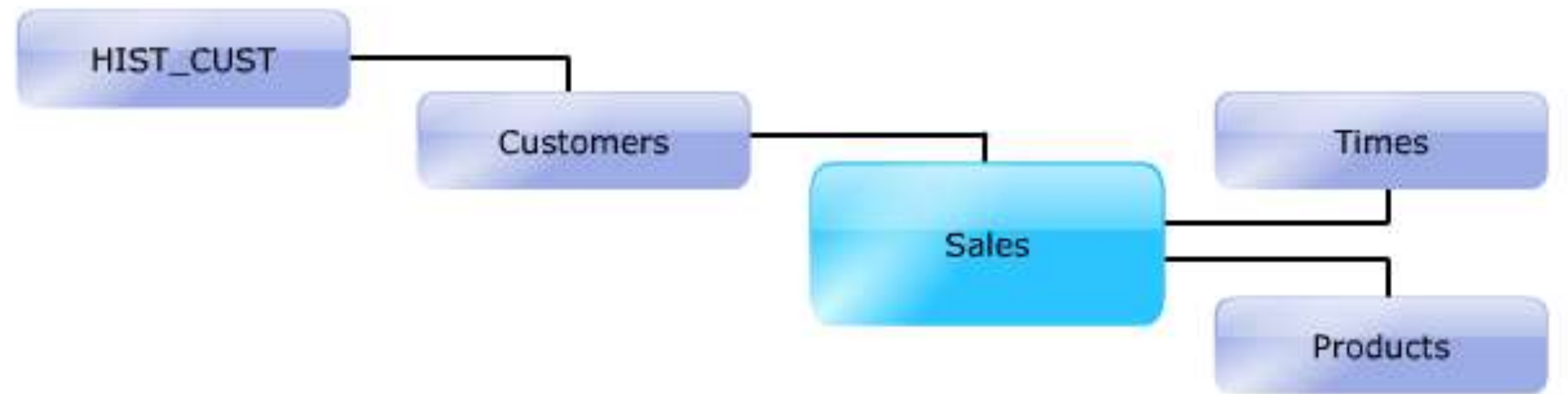


Optimize DML
during refresh





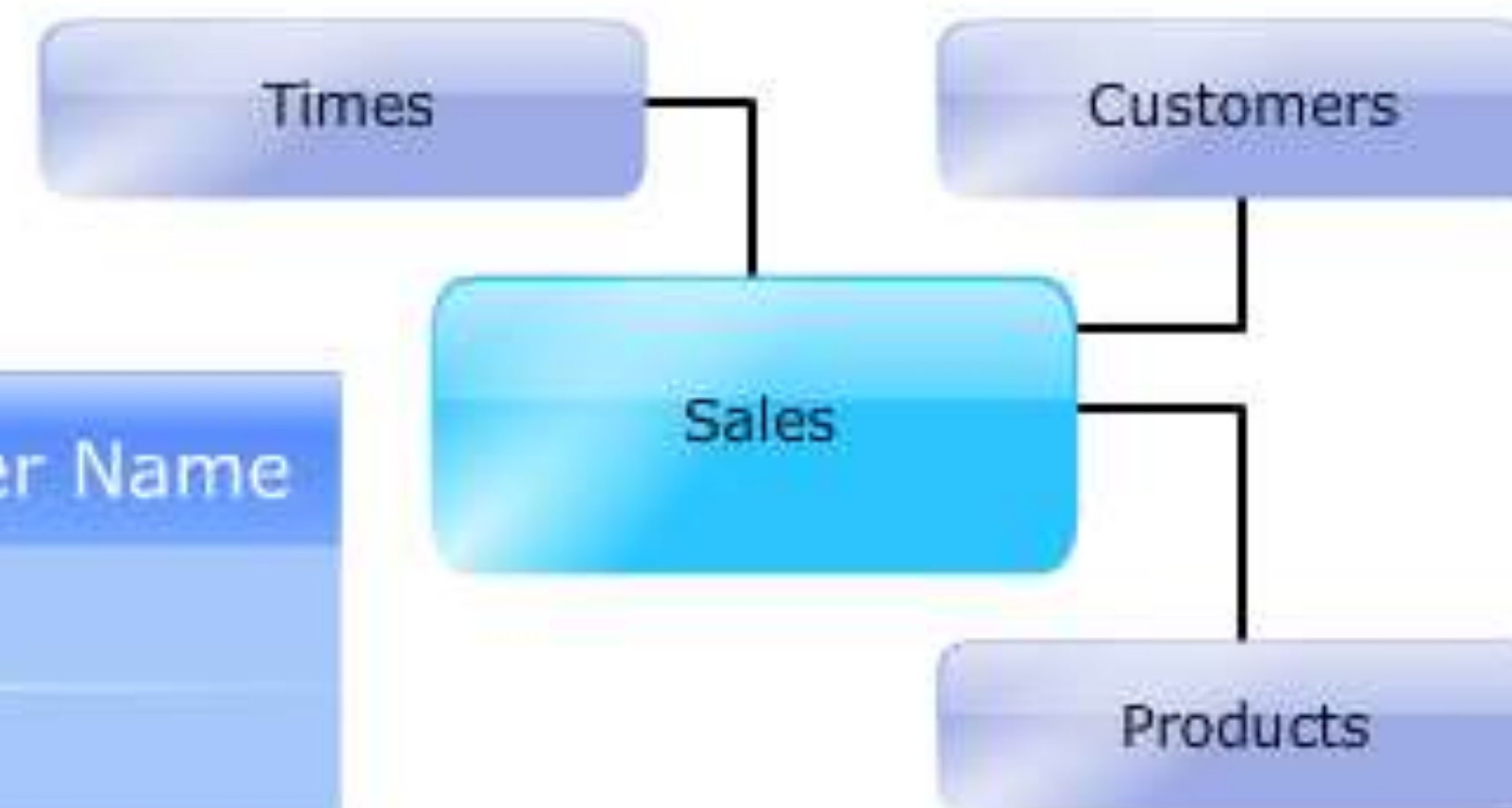
Podemos criar registros históricos em
Dimensões de dados históricos.





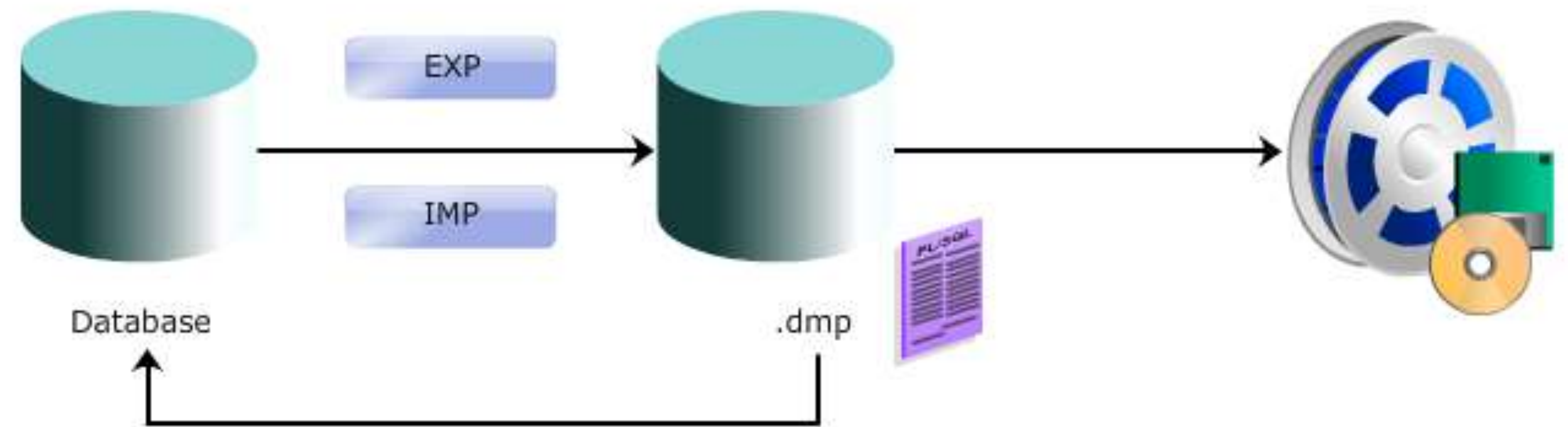
Existem várias formas de atualizar os dados no DW e podemos usar versionamento.

Customer.CustId	Version	Customer Name
1234	1	Comer
1234	2	Comer
Sales.CustId	Version	Sales Facts
1234	1	\$11,000
1234	2	\$12,000





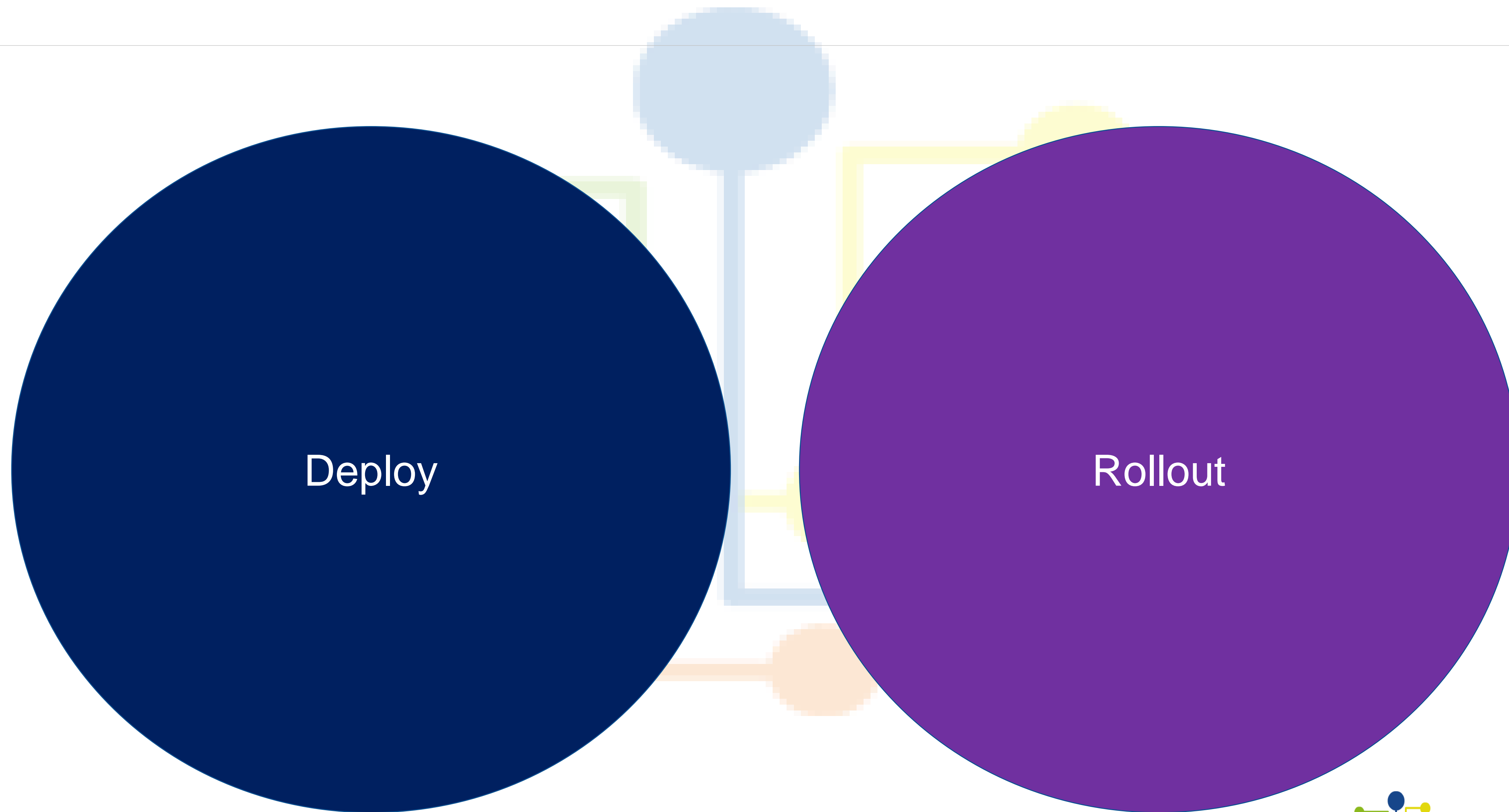
Precisamos ainda definir as estratégias de Archiving (limpeza de dados do DW). O Oracle oferece a possibilidade de fazer um dump de tabelas com o DataPump.





Deploy e Rollout







Deploy
(Colocar o DW em
produção)

Rollout
(Passar a administração
para o suporte)





Lições Aprendidas





Sessão de Lições Aprendidas

- Projeto concluído no prazo?
- Projeto concluído no orçamento?
- Entregas x Equipe
- Métodos utilizados
- Pontos Positivos x Pontos Negativos
- Principais dificuldades
- O que pode ser melhorado?
- Base de conhecimento



Data Science Academy

10 Erros Comuns a Se Evitar em Projetos de Data Warehouse





- 1- Focar em “Real-Time” ao invés de “Right Time”.
- 2- Usar a mesma infraestrutura de ETL para todos os projetos de DW.
- 3- Sumarizar demais os dados.
- 4- Não implementar Alta Disponibilidade.
- 5- Subestimar a importância dos dados históricos.
- 6- Não trabalhar na integração dos dados.
- 7- Projetar o DW apenas para as necessidades atuais.
- 8- Não priorizar a velocidade.
- 9- Considerar o DW uma solução de TI.
- 10- Não focar na qualidade dos dados.



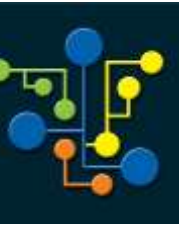


Data Science
Academy

Data Science Academy eng.davidborges@gmail.com 59532d8f5e4cdead748b456a

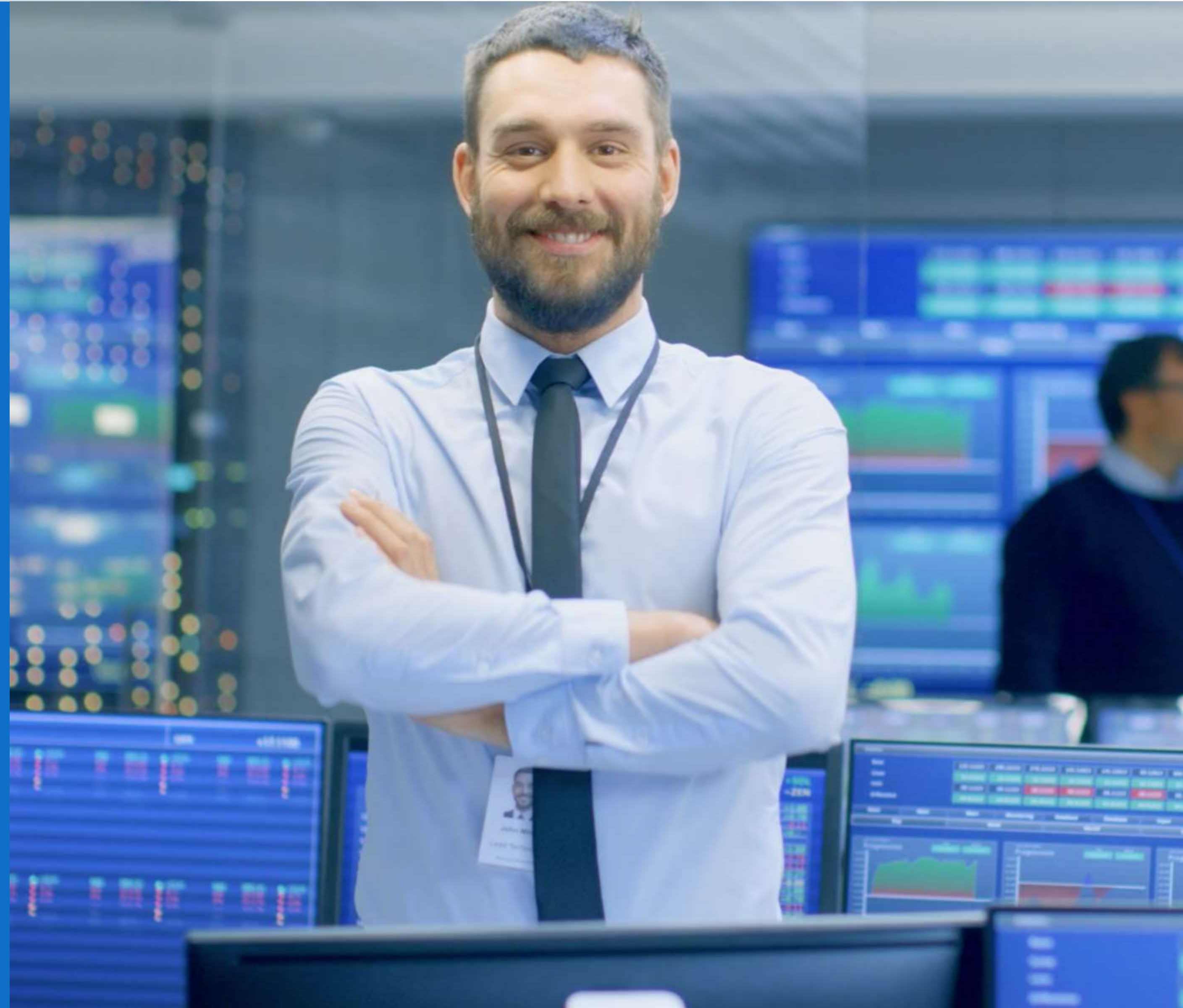
www.datascienceacademy.com.br

Design e Implementação de Data Warehouses



Data Science Academy

Métricas e KPIs





Métrica

KPI
(Key Performance Indicator)





Data Science Academy

Reporting e Segurança





Reporting e Segurança





Muito Obrigado!

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



Oportunidade



Disponibilidade



Conhecimento