

Redshift

- ◊ Solução de DW na Nuvem: AWS
- ◊ Para processar grandes volumes de dados!
- ◊ Orientado a Coluna
- ◊ Baseado em versão do Postgres
 - ◊ Possui semelhança em sintaxe SQL, mas é um produto completamente diferente
- ◊ Não tem a maioria das restrições de integridade referencial*, como chaves estrangeiras e unique



Amazon Redshift

Orientado a Coluna



Amazon Redshift

- ◇ Colunas são armazenadas individualmente
- ◇ Valores são armazenados de forma contínua
- ◇ Criar uma linha requer recuperar valores de várias colunas

Orientado a Linha VS Orientado a Coluna

- ◇ Orientado a Linha
 - ◇ Ler uma coluna, requer a leitura de todas
 - ◇ Um a linha pode ter vários tipos, então a compressão é menor
 - ◇ Ler todas as colunas tem um custo menor
 - ◇ Inserir ou atualizar colunas tem um custo menor
- ◇ Orientado a Coluna
 - ◇ Ler uma coluna, requer a leitura apenas da coluna
 - ◇ A coluna tem um mesmo tipo, então tem maior compressão
 - ◇ Ler todas as colunas tem um custo alto
 - ◇ Inserir ou atualizar colunas tem um custo maior

Ler colunas A e D, todas as linhas

A	B	C	D
123	01/01/01	20.5	Lorem ipsum
123	01/01/01	20.5	Lorem ipsum
123	01/01/01	20.5	Lorem ipsum
123	01/01/01	20.5	Lorem ipsum
123	01/01/01	20.5	Lorem ipsum
123	01/01/01	20.5	Lorem ipsum
123	01/01/01	20.5	Lorem ipsum
1 GB	1 GB	1 GB	1 GB

Orientado a Coluna:
2 GB
Orientado a Linha
4 GB