

- O modelo dimensional baseia-se na noção de _____.

O modelo dimensional baseia-se na noção de simplicidade e desempenho, organizando dados em fatos e dimensões para facilitar análises.

- Comente a seguinte afirmação:

A estrutura desnormalizada das tabelas de dimensão não proporciona a coexistência de categorias e subcategorias no mesmo espaço físico.

A estrutura desnormalizada permite categorias e subcategorias, mas pode exigir atributos redundantes para suportar diferentes níveis hierárquicos.

- Considere uma hierarquia de formato variável.

Qual é a melhor abordagem para construir essa hierarquia no modelo dimensional?

A abordagem mais eficiente é usar uma única tabela dimensional com colunas para armazenar diferentes níveis hierárquicos.

- Considere uma dimensão horária (pode denominá-la de Hora).

Dê um exemplo de uma surrogate key para essa dimensão que resolva o caso em que ainda não aconteceu uma marca horária.

Uma surrogate key pode ser um identificador numérico, como -1 para representar horários ainda não ocorridos.

- O que diferencia as chaves artificiais da dimensão DATA de todas as outras dimensões?

As chaves artificiais da dimensão DATA geralmente seguem o formato

YYYYMMDD, enquanto outras dimensões usam identificadores numéricos arbitrários.

- O que é o snowflaking?

Snowflaking é a normalização de dimensões, separando atributos em tabelas menores para reduzir redundância, mas complicando consultas.

- Qual é o propósito das hierarquias nas tabelas de factos?

Hierarquias nas tabelas de fatos permitem análises em diferentes níveis de agregação, como tempo, produto ou localização.

- Qual o número máximo de tabelas a que pode estar ligada uma dimensão degenerada?

Uma dimensão degenerada pode estar ligada a várias tabelas de fatos, sem um limite específico, dependendo do modelo de dados.

- Considere os processos de negócio.

Qual o principal erro que se pode fazer no desenho de um data mart?

O principal erro é não considerar o modelo de negócio corretamente, resultando em um data mart ineficaz para análise.

O principal erro é não alinhar o data mart às necessidades do negócio, resultando em dados irrelevantes ou incompletos.

- Os dados no DW contêm uma _____ de tal forma que podem ser utilizados para estudar, por exemplo, _____.

Os dados no DW contêm estrutura histórica para estudar tendências, padrões de consumo, desempenho e outras análises estratégicas.

- Os interesses dos _____ são os únicos que têm sempre que ser defendidos no desenvolvimento do data warehouse.

Os interesses dos usuários são prioritários no desenvolvimento do data warehouse,

garantindo utilidade e acessibilidade dos dados.

12. Um data warehouse contém _____ que descrevem a forma como os dados estão organizados.

Um data warehouse contém metadados, descrevendo estrutura, relacionamento e significado dos dados armazenados.

13. Das opções seguintes indique todas as que apresentam exemplos de hierarquias.

Exemplos de hierarquias incluem tempo (ano, mês, dia), geografia (país, estado, cidade) e produto (categoria, subcategoria, item).

14. As tabelas de acumulação periódica servem para:

As tabelas de acumulação periódica armazenam fatos agregados ao longo do tempo para análises históricas e comparativas.

15. O que é uma tabela de mini-dimensão?

Uma mini-dimensão armazena atributos frequentemente alterados, reduzindo impactos nas dimensões principais e facilitando a análise.

16. Considere o valor unitário de uma embalagem de aspirina com 20 comprimidos.

Esse valor será um facto de que tipo?

O valor unitário é um fato semi-aditivo, pois pode ser agregado apenas em algumas dimensões.

17. As tabelas de factos de acumulação periódica diária são as que ocupam mais espaço no data warehouse.

Verdadeiro ou Falso?

Verdadeiro, pois tabelas de factos diárias acumulam grande volume de dados, exigindo mais espaço.

18. Qual é o principal benefício da abordagem de Inmon, no desenho do DW?

Inmon propõe um DW centralizado e normalizado, garantindo flexibilidade, consistência e integração entre diferentes áreas da empresa.

19. Uma hierarquia no modelo dimensional é:

Uma hierarquia no modelo dimensional representa relações entre níveis de dados, como dia-mês-ano ou produto-categoria.

20. Indique as propriedades que são exigidas para que uma dimensão seja considerada conforme.

Para ser conforme, uma dimensão deve ser padronizada, consistente e reutilizável em múltiplos data marts.

21. Habitualmente, a quantas fontes de dados pode ser associado um processo de negócio?

Um processo de negócio pode ser associado a várias fontes de dados, dependendo da complexidade operacional.

22. Os factos têm uma propriedade interessante que é o de poderem ser _____ segundo um critério.

Os fatos podem ser agregados segundo um critério, permitindo análises detalhadas ou resumidas.

23. A maior parte dos data warehouses usa sistemas de bases de dados _____ para controlar e gerir a informação.

A maioria dos data warehouses usa sistemas de bases de dados relacionais para armazenamento e gerenciamento de informações.

24. O que é um Data Warehouse?

Um Data Warehouse é um repositório centralizado para análise e tomada de decisões baseadas em dados históricos.

25. Qual a principal função de uma tabela de dimensão?

A principal função de uma tabela de dimensão é descrever o contexto dos fatos armazenados.

26. O que contém uma tabela de fato?

Uma tabela de fato contém métricas numéricas que representam eventos de negócios, referenciando dimensões.

27. O que é o "grão" num modelo dimensional?

O "grão" define o nível de detalhe dos fatos, como vendas por dia ou por transação.

28. Qual o tipo de fato que só pode ser agregado em certas dimensões?

Fatos semi-aditivos só podem ser agregados em certas dimensões, como saldo bancário ao longo do tempo.

29. O que é um esquema em estrela?

Um esquema em estrela tem tabelas de fatos conectadas diretamente a dimensões desnormalizadas, otimizando consultas.

30. Para que serve uma mini dimensão?

Uma mini dimensão armazena atributos voláteis para evitar alterações frequentes em dimensões principais.

31. O que é uma tabela outrigger?

Uma tabela outrigger é uma dimensão normalizada separada para atributos raramente usados.

32. Qual o grão de uma tabela de facturação com dados por dia e produto?

O grão é vendas diárias por produto, ou seja, cada linha representa um produto vendido em determinado dia.

33. Qual o tipo de chave mais comum em dimensões?

A chave mais comum em dimensões é a surrogate key, um identificador numérico sem significado real.

34. O que é a dimensão "Tempo"?

A dimensão "Tempo" organiza dados temporais, permitindo análises sazonais, tendências e previsões.

35. Qual o propósito de uma tabela de fato de acumulação?

Uma tabela de fato de acumulação registra métricas agregadas ao longo do tempo, facilitando análises históricas.

36. O que caracteriza uma dimensão de alteração lenta (Slowly Changing Dimension)?

Uma dimensão de alteração lenta (SCD) armazena mudanças históricas de atributos sem sobrescrever dados antigos.

37. Qual o tipo de fato que não pode ser agregado?

Fatos não aditivos, como taxas percentuais, não podem ser somados diretamente.

38. O que é um esquema floco de neve (snowflake)?

O esquema floco de neve normaliza dimensões para reduzir redundância, mas dificulta consultas.

39. Para que serve uma tabela bridge?

Uma tabela bridge resolve relações muitos-para-muitos entre dimensões e fatos.

40. Qual o tipo de grão mais granular?

O tipo de grão mais granular é aquele que registra dados no menor nível possível, como transações individuais.

41. O que é uma dimensão degenerada?

Uma dimensão degenerada armazena atributos diretamente na tabela de fatos, sem tabela separada.

42. Qual a vantagem do esquema em estrela?

O esquema em estrela simplifica consultas e melhora o desempenho, pois reduz a necessidade de joins complexos.

43. O que é uma hierarquia de dimensão?

Uma hierarquia de dimensão organiza dados em diferentes níveis de granularidade, como país → estado → cidade.

44. Qual a função de um atributo derivado?

Um atributo derivado é calculado com base em outros atributos, otimizando consultas e reduzindo redundância.

45. Qual o tipo de facto que representa custos fixos?

Fatos de custos fixos não variam com a quantidade de produtos vendidos, sendo alocados proporcionalmente.

46. O que é uma tabela de facto de snapshot?

Uma tabela de fato de snapshot registra o estado de métricas em momentos específicos.

47. Qual a desvantagem do esquema floco de neve?

O esquema floco de neve complica consultas devido a joins adicionais entre tabelas normalizadas.

48. O que é uma tabela de facto de transição?

Uma tabela de fato de transição rastreia eventos intermediários em processos de negócios.

49. Qual a finalidade de uma dimensão de data?

A dimensão de data permite análises temporais estruturadas no data warehouse.

50. O que é uma junk dimension?

Uma junk dimension agrupa atributos irrelevantes em uma única tabela para reduzir a complexidade dimensional.

51. Qual o tipo de chave usada para ligar factos a dimensões?

A chave mais usada para ligar fatos a dimensões é a surrogate key, um identificador numérico gerado.

52. Qual a principal diferença entre OLTP e OLAP?

OLTP processa transações operacionais diárias, enquanto OLAP é otimizado para análise e relatórios estratégicos.