

A velocidade é uma característica importante em projetos Big Data uma vez que com o aumento da velocidade dos dispositivos conectados à internet e das tecnologias móveis, os dados estão sendo gerados em uma taxa cada vez maior. Isso significa que as empresas precisam ser capazes de coletar, processar e analisar grandes volumes de dados em tempo real para obter insights valiosos e tomar decisões informadas. Além disso, a velocidade também é importante para garantir que as empresas possam responder rapidamente a eventos em tempo real, como fraudes ou problemas de segurança. Portanto, a capacidade de lidar com grandes volumes de dados em alta velocidade é fundamental para o sucesso dos projetos Big Data.

BASSO, D. E. Big Data. Curitiba: Contentus, 2020. Pag 29-33. Adaptado.

Considerando o texto anterior, avalie as afirmações a seguir sobre a relação da velocidade e grandes volumes de dados.

- I. A velocidade pode ser resolvida com o uso de sistemas centralizados e bancos de dados locais.
- II. A velocidade lida com a eficiência do processamento de dados. Levando a atrasos na entrega dos resultados da análise.
- III. A velocidade pode afetar a qualidade dos dados, como duplicação de dados, perda de dados ou inconsistências nos dados.
- IV. A velocidade pode levar a problemas de escalabilidade e aumento de custos, pela demanda de aumento de armazenamento.

É correto o que se afirma em:

- ( ) A II, III e IV, apenas.
- ( ) B III e IV, apenas.
- ( ) C I e IV, apenas.
- (x) D II e III, apenas.
- ( ) E I, II e III, apenas.

## QUESTÃO 2 | 0,40

---

A análise de séries temporais permite identificar tendências, padrões sazonais e anomalias ao longo do tempo, sendo fundamental para a compreensão de dados históricos e de projeções futuras. Em projetos complexos, esse tipo de análise pode informar decisões estratégicas em diferentes níveis organizacionais, desde a operação até a gestão. A correta interpretação dessas informações contribui para o planejamento eficiente e para a otimização de processos. Assim, compreender quais níveis são beneficiados torna-se essencial para direcionar esforços e maximizar resultados.

Com base no texto apresentado, pela análise de séries temporais no projeto, quais são os principais níveis beneficiados?

---

- ( )  A Apenas o operacional.
- ( )  B Financeiro e administrativo.
- (x)  C Operacional e gerencial.
- ( )  D Clínico e jurídico.
- ( )  E Apenas o estratégico.

## QUESTÃO 3 | 0,40

---

Bancos de dados relacionais e não relacionais diferem fundamentalmente em sua estrutura e abordagem para o armazenamento de dados. Bancos de dados relacionais, como MySQL e PostgreSQL, organizam dados em tabelas e utilizam SQL para gerenciar as relações entre essas tabelas. Em contraste, bancos de dados não relacionais (NoSQL), como MongoDB e Neo4j, oferecem flexibilidade ao armazenar dados em formatos variados, como documentos, colunas, pares chave-valor ou grafos (Saravis, 2023).

Levando em consideração os bancos de dados baseados em tabelas, como o MySQL, e aqueles baseados em documentos, como o MongoDB, analise as sentenças a seguir:

- I. Bancos de dados relacionais requerem manutenção de índices de forma manual, enquanto bancos de dados baseados em documentos geram índices automaticamente para todos os campos.
- II. Em bancos de dados relacionais, a integridade dos dados é mantida através de relações e chaves estrangeiras, enquanto em bancos de dados baseados em documentos, a integridade é mantida através de referências e embeddings.
- III. Bancos de dados relacionais são otimizados para operações de leitura e escrita rápidas, enquanto bancos de dados baseados em documentos são otimizados apenas para operações de leitura.
- IV. Bancos de dados baseados em documentos permitem uma modelagem de dados mais natural para aplicações com dados hierárquicos ou complexos, enquanto bancos de dados relacionais são mais adequados para dados estruturados e normalizados.
- V. Bancos de dados relacionais armazenam dados em tabelas com esquemas rígidos, enquanto bancos de dados baseados em documentos armazenam dados em documentos JSON-like com esquemas flexíveis.

É correto o que se afirma em:

- ( ) A I, III e IV, apenas.
- (x) B II, IV e V, apenas.
- ( ) C I, II e V, apenas.
- ( ) D II, III e IV, apenas.
- ( ) E II, III e V, apenas.

#### QUESTÃO 4 | 0,40

Uma empresa de logística global está implementando um sistema de rastreamento de mercadorias que precisa armazenar dados de sensores IoT instalados em contêineres de carga. Os sensores coletam informações sobre temperatura, umidade, localização GPS, vibração e abertura de portas durante o transporte marítimo e rodoviário. O sistema deve permitir consultas históricas sobre condições de transporte e detectar automaticamente violações de temperatura para cargas sensíveis, como medicamentos e alimentos.

Considerando as informações apresentadas, avalie as asserções a seguir e a relação proposta entre elas:

I. Dados de sensores IoT com coleta temporal frequente são melhor gerenciados por SGBDs de séries temporais do que por bancos relacionais tradicionais.

PORQUE

II. Bancos de séries temporais oferecem compressão otimizada para dados temporais e funções específicas para análise de tendências e detecção de anomalias.

A respeito dessas asserções, assinale a alternativa correta:

- ( ) A A asserção I é uma proposição verdadeira, e a II é uma proposição falsa.
- ( ) B A asserção I é uma proposição falsa, e a II é uma proposição verdadeira.
- ( ) C As asserções I e II são proposições verdadeiras, mas a II não é uma justificativa correta da I.
- ( ) D As asserções I e II são proposições falsas.
- (x) E As asserções I e II são proposições verdadeiras, e a II é uma justificativa correta da I.

Uma empresa de logística global está implementando um sistema de rastreamento básico que coleta dados simples de localização GPS dos contêineres apenas para relatórios de conformidade regulatória. O sistema não precisa de análise temporal complexa, apenas armazenamento e consultas por período. A equipe técnica está debatendo se realmente precisam de SGBDs especializados em séries temporais ou se podem usar bancos relacionais tradicionais com campos de data/hora.

Considerando as informações apresentadas, avalie as asserções a seguir e a relação proposta entre elas.

I. SGBDs de séries temporais são sempre obrigatórios para qualquer dado que contenha *timestamps*, mesmo para uso simples.

PORQUE

II. Dados temporais sempre exigem funcionalidades avançadas, como agregação temporal e detecção automática de anomalias.

A respeito dessas asserções, assinale a alternativa correta:

- ( x )  A As asserções I e II são proposições falsas.
- ( )  B As asserções I e II são proposições verdadeiras, mas a II não é uma justificativa correta da I.
- ( )  C A asserção I é uma proposição verdadeira, e a II é uma proposição falsa.
- ( )  D As asserções I e II são proposições verdadeiras, e a II é uma justificativa correta da I.
- ( )  E A asserção I é uma proposição falsa, e a II é uma proposição verdadeira.

## QUESTÃO 6 | 0,40

No contexto de Bancos de Dados baseados em documentos, como o MongoDB, diferentes formatos de dados como JSON (JavaScript Object Notation), BSON (Binary JSON), YAML (YAML Ain't Markup Language) e XML (eXtensible Markup Language) podem ser utilizados para armazenar e transmitir dados.

Fonte: SAVARIS, A. **Banco de dados NoSQL**. Maringá: UniCesumar, 2022.

Considerando esses formatos de dados, analise as afirmativas a seguir:

- I. YAML (YAML Ain't Markup Language) é um formato de serialização de dados que é mais legível para humanos e frequentemente usado para arquivos de configuração, mas não é adequado para armazenamento em bancos de dados devido à sua complexidade.
- II. JSON (JavaScript Object Notation) é um formato leve de intercâmbio de dados que é fácil para humanos lerem e escreverem, e fácil para máquinas interpretarem e gerarem.
- III. BSON (Binary JSON) é uma versão textual do JSON que permite a inclusão de tipos de dados adicionais como datas e binários, além de ser mais eficiente em termos de espaço e velocidade de leitura/escrita.
- IV. XML (eXtensible Markup Language) é um formato de dados autodescritivo que usa tags para definir objetos e dados, mas é menos eficiente que JSON e BSON em termos de espaço e processamento.
- V. Em bancos de dados baseados em documentos, como o MongoDB, o BSON é o formato de dados nativo utilizado para armazenar documentos internamente devido à sua eficiência e capacidade de suporte a tipos de dados adicionais.

É correto o que se afirma em:

- ( )  A I, III e V, apenas.
- ( )  B I, IV e V, apenas.
- (x)  C III, IV e V, apenas.
- ( )  D II, IV e V, apenas.
- ( )  E II, III e V, apenas.

## QUESTÃO 7 | 0,40

Uma empresa de *streaming* de vídeo está enfrentando problemas de performance durante eventos esportivos transmitidos ao vivo, onde milhões de usuários acessam simultaneamente o conteúdo. O sistema atual baseado em PostgreSQL não consegue lidar com o pico de requisições de metadados dos vídeos (título, descrição, *tags*, tempo de duração). A equipe técnica precisa escolher uma solução NoSQL que ofereça latência ultra-baixa para operações de leitura simples, já que os metadados são raramente modificados, mas consultados milhões de vezes por minuto durante eventos populares.

Com base no texto apresentado, para esse caso específico de alta performance em leituras simples, qual modelo NoSQL oferece melhor adequação?

- ( )  A Séries temporais, pelo aspecto temporal dos eventos.
- ( )  B Orientado a documentos, pela flexibilidade de metadados.
- ( )  C Orientado a grafos, pelos relacionamentos entre conteúdos.
- (x)  D Chave-valor, pela simplicidade e alta performance.
- ( )  E Orientado a colunas, pela compressão de dados.

## QUESTÃO 8 | 0,40

A teoria dos grafos é uma ferramenta matemática poderosa para modelar e analisar relações entre objetos ou entidades em diversas áreas do conhecimento. A estrutura de um grafo é composta por nós e arestas. Os nós são as unidades básicas que representam os objetos ou entidades do conjunto em questão. Já as arestas representam as relações entre esses objetos ou entidades. As arestas podem ser direcionadas ou não, o que significa que elas podem ter uma direção definida ou não. A compreensão dessa estrutura básica é fundamental para entender as aplicações práticas dos grafos e desenvolver algoritmos eficientes para resolução de problemas complexos.

NETTO, P. O. B; JURKIEWICZ, S. Grafos: introdução e prática. São Paulo: Blucher, 2009. Adaptado.

Com base nas informações apresentadas, avalie as asserções a seguir e a relação proposta entre elas:

I. A estrutura de um grafo é composta por nós e arestas, que representam objetos ou entidades e suas relações, respectivamente.

PORQUE

II. A classificação de um grafo como direcionado ou não direcionado depende da existência ou não de um peso atribuído às arestas.

A respeito dessas asserções, assinale a opção correta:

- ( )  A As asserções I e II são falsas.
- ( )  B A asserção I é uma proposição falsa e a II é uma proposição verdadeira.
- (x)  C **A asserção I é uma proposição verdadeira e a II é uma proposição falsa.**
- ( )  D As asserções I e II são verdadeiras, mas a II não é uma justificativa correta da I.
- ( )  E As asserções I e II são verdadeiras, e a II é uma justificativa correta da I.

Séries temporais são sequências de dados coletados ao longo do tempo, frequentemente em intervalos regulares. Elas permitem identificar padrões, tendências e sazonalidades em fenômenos variados, como finanças, clima ou redes de computadores. O estudo dessas séries é essencial para previsões e tomadas de decisão baseadas em análise histórica.

Considerando o texto apresentado, sobre a implementação das séries temporais no projeto, analise as afirmativas a seguir:

- I. A metodologia de análise de séries temporais já possui respaldo teórico.
- II. A decisão de usar séries temporais permite decisões baseadas em dados.
- III. O projeto não considera o nível gerencial para tomada de decisão.
- IV. Série temporal só existe em banco relacional.

É correto o que se afirma em:

- ( )  A III e IV, apenas.
- ( )  B I, II, III e IV.
- ( )  C II e IV, apenas.
- ( )  D I e III, apenas.
- (x)  E I e II, apenas.

Uma série temporal é uma sequência de observações de uma variável ao longo do tempo, que pode ser utilizada para identificar padrões e tendências ao longo do período observado. A análise de tendências em uma série temporal é um importante aspecto da análise de dados, pois permite que sejam identificadas mudanças significativas na variável ao longo do tempo e que sejam feitas previsões para o futuro. Existem diversas técnicas para a análise de tendências em séries temporais, sendo uma das mais comuns a análise de regressão, a análise de médias móveis, a decomposição da série temporal em componentes sazonais e a análise de séries temporais em espaço de estado.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 3. ed. São Paulo: Blucher, 2018. Adaptado.

Considerando o texto anterior, assinale a alternativa que indica o comportamento que descreve uma tendência em uma série temporal.

- ( )  A Comportamento aleatório dos dados ao longo do tempo.
- ( )  B Comportamento linear dos dados ao longo do tempo.
- (x)  C Comportamento periódico dos dados ao longo do tempo.
- ( )  D Comportamento de longo prazo que pode ser observado nos dados.
- ( )  E Comportamento sazonal (diário ou semanal, por exemplo).