

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO INSTITUTO NACIONAL DE PESQUISA DO PANTANAL (INPP)

CARGO 7: TECNOLOGISTA PLENO 1 ESPECIALIDADE: ANALISTA DE BANCO DE DADOS

Prova Discursiva

Aplicação: 02/02/2025

PADRÃO DE RESPOSTA DEFINITIVO

TEMA 10

Recursos de SQL para o apoio às atividades de análise de dados

O(A) candidato(a) deve apresentar uma discussão acerca dos principais recursos e das funcionalidades da linguagem SQL que facilitem a extração, a manipulação, a análise e a garantia da qualidade dos dados. Espera-se que o(a) candidato(a):

1. **Fundamente o papel da SQL na análise de dados**, explicando sua importância na obtenção de informações relevantes para a tomada de decisões, na integração com sistemas de Business Intelligence (BI) e na manutenção da qualidade e integridade dos dados;
2. **Descreva os recursos básicos e avançados da SQL**, abordando desde os comandos essenciais e a estrutura de uma consulta até técnicas mais sofisticadas, tais como funções agregadas, funções de janela, subconsultas, operações de junção e expressões regulares/manejo de caracteres, enfatizando os comandos de manipulação (DML) e a definição (DDL);
3. **Apresente um estudo de caso ou exemplo prático** que demonstre como os recursos da SQL podem ser aplicados para diagnosticar problemas ou identificar tendências em conjuntos de dados, ressaltando a integração com ferramentas de análise, a proposição de soluções e a garantia da qualidade dos dados, inclusive através de técnicas de manipulação de caracteres e expressões regulares para limpeza e validação dos dados.

QUESITO 2.1 – Fundamentos e conceitos básicos da SQL

(i) Definição e importância da SQL:

- O(A) candidato(a) deve contextualizar o papel da SQL no ambiente de bancos de dados e na análise de dados, destacando sua relevância para a extração de informações e para a garantia da qualidade e da integridade dos dados.

(ii) Principais comandos e estruturas básicas:

- Espera-se a menção aos comandos fundamentais (por exemplo, SELECT, INSERT, UPDATE, DELETE) e a explicação da estrutura básica de uma consulta, abordando cláusulas essenciais (FROM, WHERE, etc.).

(iii) Classificação dos comandos SQL – ênfase em DML e DDL:

- Deve ser explicado que, para as atividades de análise de dados, os comandos de manipulação (DML) e de definição (DDL) são os mais relevantes, deixando de lado a DCL (voltada para controle de acesso e segurança).

(iv) Aspectos relativos à qualidade dos dados:

- O(A) candidato(a) deve mencionar como a SQL contribui para a verificação, a validação e a manutenção da qualidade dos dados, por meio de mecanismos de integridade, validação e limpeza.

QUESITO 2.2 – Recursos avançados da SQL para análise de dados

(i) Uso de funções agregadas e de janela:

- O(A) candidato(a) deve mencionar funções como SUM, COUNT, AVG, MAX, MIN e explicar o uso de funções de janela (ex.: OVER, PARTITION BY) para análises mais detalhadas e para a identificação de padrões ou inconsistências que possam influenciar a qualidade dos dados.

(ii) Operações de junção e subconsultas:

- Espera-se a explicação das diferentes junções (INNER JOIN, LEFT JOIN, etc.) e o emprego de subconsultas para combinar e filtrar conjuntos de dados complexos, ressaltando como essas operações auxiliam na obtenção de dados coerentes e na melhoria da qualidade da informação.

(iii) Expressões regulares e manipulação de caracteres:

- Deve ser abordada a utilização de expressões regulares e funções de manipulação de caracteres (como funções para busca, substituição ou extração de padrões) na SQL, evidenciando como essas técnicas contribuem para a validação, limpeza e transformação dos dados.

(iv) Utilização de cláusulas avançadas:

- Devem ser mencionadas cláusulas como GROUP BY, HAVING e ORDER BY, que auxiliam na organização, na agregação e na interpretação dos dados.

QUESITO 2.3 – Estudo de caso ou exemplo de aplicação prática da SQL na análise de dados

(i) Contextualização do problema:

- O(A) candidato(a) deve apresentar um cenário prático – por exemplo, a análise do desempenho de vendas, o comportamento de clientes ou o monitoramento de indicadores operacionais – que justifique a necessidade do uso dos recursos avançados da SQL, com atenção especial à obtenção e à garantia da qualidade dos dados.

(ii) Uso das ferramentas da SQL no diagnóstico das questões:

- Espera-se a descrição de como os comandos e as funções da SQL foram empregados para extrair e manipular dados relevantes, contribuindo para a identificação de problemas, inconsistências ou padrões que possam afetar a qualidade dos dados, incluindo o uso de expressões regulares e funções de manipulação de caracteres, quando aplicável.

(iii) Proposição de soluções e integração com ferramentas de análise:

- O(A) candidato(a) deve demonstrar como os resultados obtidos com a SQL podem ser integrados a ferramentas de BI (como *dashboards* e relatórios) para a tomada de decisão e a proposição de soluções, ressaltando a importância de manter a qualidade dos dados para análises precisas.

QUESITOS AVALIADOS

QUESITO 2.1

Conceito 0 – Não desenvolveu o quesito ou o fez de forma totalmente equivocada.

Conceito 1 – Discorreu de forma superficial sobre a definição da SQL, sem abordar os demais itens.

Conceito 2 – Abordou a SQL e apresentou apenas um dos itens adicionais (por exemplo, apenas a estrutura básica de consulta).

Conceito 3 – Desenvolveu dois dos itens especificados (por exemplo, definição e estrutura básica dos comandos), sem aprofundar os demais pontos.

Conceito 4 – Discorreu de forma completa, abordando os quatro itens especificados, inclusive os aspectos relativos à qualidade dos dados.

QUESITO 2.2

Conceito 0 – Não desenvolveu o quesito ou o fez de forma totalmente equivocada.

Conceito 1 – Descreveu adequadamente apenas um dos itens especificados.

Conceito 2 – Descreveu dois dos itens especificados, mas cometeu algum erro conceitual ou deixou de integrar completamente as ideias.

Conceito 3 – Descreveu corretamente três dos itens especificados, integrando o uso dos recursos avançados da SQL (inclusive expressões regulares e manipulação de caracteres).

Conceito 4 – Descreveu de forma completa e integrada todos os itens especificados, demonstrando domínio dos recursos avançados e sua aplicação na melhoria da qualidade dos dados.

QUESITO 2.3

Conceito 0 – Não desenvolveu o quesito ou o fez de forma totalmente equivocada.

Conceito 1 – Desenvolveu adequadamente apenas um dos itens especificados (por exemplo, apresentou o problema sem explicar a aplicação prática da SQL).

Conceito 2 – Desenvolveu adequadamente dois dos itens especificados, mas deixou de integrar ou aprofundar a solução prática.

Conceito 3 – Desenvolveu os três itens especificados, embora possa ter cometido algum erro conceitual ou deixado algum aspecto de forma menos aprofundada.

Conceito 4 – Desenvolveu de forma completa e integrada os três itens especificados, apresentando um estudo de caso consistente que evidencia o uso dos recursos da SQL no diagnóstico, na proposição de soluções e na garantia da qualidade dos dados analisados.