**Métricas de Observabilidade para o Processo ETL do sales\_data.csv:**

1. **Tempo de Extração, Transformação e Carregamento (ETL):**
   * Tempo total gasto em cada etapa (extração, transformação e carregamento) do ETL.
   * Tempo médio de processamento por registro em cada etapa.
   * Alertas configurados para alertar quando o tempo de processamento excede limites definidos.
2. **Monitoramento de Dados:**
   * Número total de registros extraídos do arquivo CSV.
   * Número total de registros após a transformação.
   * Número total de registros carregados no banco de dados.
   * Monitoramento da integridade dos dados, como verificar se todos os campos obrigatórios estão presentes e têm valores válidos.
3. **Monitoramento de Erros e Anomalias:**
   * Número de registros rejeitados durante a transformação devido a problemas de qualidade dos dados.
   * Número de erros de carga no banco de dados.
   * Monitoramento de exceções e erros no pipeline ETL.
   * Alertas configurados para notificar quando o número de erros exceder um limite aceitável.
4. **Monitoramento de Desempenho de Banco de Dados:**
   * Uso de recursos do banco de dados (CPU, memória, espaço em disco) durante a carga.
   * Taxa de gravação no banco de dados durante a carga.
   * Alertas para indicar problemas de desempenho no banco de dados.
5. **Monitoramento de Integridade de Dados:**
   * Verificação de duplicatas ou valores ausentes em campos únicos.
   * Monitoramento de chaves estrangeiras para garantir integridade referencial.

**Implementação de Alertas:**

1. **Alertas em Tempo Real:**
   * Configuração de alertas em ferramentas de monitoramento como **Amazon CloudWatch** ou **Prometheus**.
   * Alertas acionados quando métricas ultrapassam limites predefinidos, como tempo de processamento muito longo ou número excessivo de erros.
2. **Alertas de Erros e Exceções:**
   * Uso de logging estruturado para registrar exceções e erros detalhados.
   * Alertas configurados para monitorar logs e acionar alertas quando mensagens de erro específicas são registradas.

**Rastreamento de Problemas:**

1. **Alertas Recebidos:**
   * Quando um alerta é recebido, a equipe é notificada via email ou via chat.
2. **Investigação Inicial:**
   * A equipe verifica o dashboard de métricas para identificar a métrica específica que causou o alerta.
   * Verifica se o problema é um evento único ou recorrente.
3. **Rastreamento no Log:**
   * Se necessário, a equipe consulta os logs detalhados do pipeline ETL para entender a causa raiz do problema.
   * Logs podem conter informações sobre exceções, erros de validação, registros rejeitados, etc.
4. **Análise do Fluxo ETL:**
   * A equipe rastreia o fluxo ETL para identificar onde o problema ocorreu (extração, transformação ou carregamento).
5. **Correção e Reteste:**
   * Uma vez identificado o problema, a equipe faz correções no pipeline ETL e retesta o processo.
   * Novas métricas são monitoradas para garantir que o problema tenha sido resolvido.