

1. Por que armazenar as respostas das APIs?

R: Existe alguns motivos para armazenar as respostas das APIs, um deles é a rastreabilidade e facilidade para auditoria futura, armazenar as respostas permite rastrear divergências futuras, validar dados com fontes originais e auditar movimentações. Outro motivo é a fonte confiável como já dito brevemente, as APIs podem mudar, sair do ar ou ser atualizadas. Ter os dados armazenados assegura consistência histórica. Também facilita análises futuras, reprocessamento de relatórios ou construção de novos pipelines sem necessidade de reconsultar as APIs.

2. Como você armazenaria os dados? Crie uma estrutura de pastas capaz de armazenar as respostas da API. Ela deve permitir manipulação, verificações, buscas e pesquisas rápidas.

R:

/data-lake/raw/cb_lab_api/

└─ [endpoint-name]/

└─ year=2025/

└─ month=07/

└─ day=25/

└─ storeId=001/

└─ response.json

Espera-se dos ENDPOINTS arquivos como: /bi/getFiscalInvoice | /res/getGuestChecks | /org/getChargeBack | /trans/getTransactions | /inv/getCashManagementDetails

3. Considere que a resposta do endpoint getGuestChecks foi alterada, por exemplo, guestChecks.taxes foi renomeado para guestChecks.taxation. O que isso implicaria?

R: Traria sim algumas implicações, provavelmente iria quebrar em pipelines que esperam o campo .taxes. Erros em parsing, transformações ou geração de relatórios. Necessidade de versionamento ou validação de schema.