

Metodologia de Coleta, Tratamento e Publicação de Dados para o Painel Fiocruz Transparente

Introdução

Documento consolidado que descreve, de forma técnica e operacional, a metodologia atual de ETL e automação utilizada no Painel Fiocruz Transparente (financeiro). Substitui e complementa as seções pertinentes do documento original, mantendo definições, siglas e referências orçamentárias.

Fontes de dados e periodicidade

- Fontes primárias: SIAFI; Limites Orçamentários da COGEPLAN (planilha mensal); Relação de Unidades Responsáveis por TEDs (SEOR/COGEAD).
- Cobertura: LOA (inclui emendas e alocação para Covid-19), Restos a Pagar e TEDs.
- Periodicidade: coleta diária; dados refletem o fechamento do dia anterior.

Ingestão e transferência automatizada

- Origem dos arquivos: Tesouro Gerencial — três arquivos em Excel fornecidos por subscrição via e-mail.
- Automação de ingestão: fluxo implementado em Power Automate que captura anexos das subscrições e grava automaticamente em pasta dedicada no OneDrive for Business.
- Validação inicial: checagens automáticas no Power Automate para confirmar presença de arquivo, formato e nomenclatura esperada antes de persistência no OneDrive.

Camada de transformação ETL

- Ferramenta: Tableau Prep.
- Pipeline: limpeza (remoção de duplicatas, tratamento de nulos), padronização de campos (códigos, datas, valores), enriquecimento e junções entre as três bases recebidas.
- Artefato: extração consolidada publicada na pasta de extrações do OneDrive.
- Qualidade: regras de validação configuradas no fluxo do Prep; discrepâncias registradas em arquivo de log.
-

Orquestração e execução agendada

- Agendador: Windows Task Scheduler.
- Mecanismo de execução: script em lote (.bat) que invoca o fluxo do Tableau Prep e gerencia o sequenciamento das tarefas.
- Horário: execução diária às 07h00.
- Logs: .bat grava saída detalhada (timestamp, status por etapa, mensagens de erro) em arquivo de log centralizado no OneDrive.
- Notificações: alertas automáticos via e-mail ou Teams acionados em falhas críticas.

Publicação do painel

- Ferramenta de visualização: Tableau Desktop.
- Procedimento operacional: abertura manual do Tableau Desktop às 08h00; atualização das extrações a partir dos arquivos consolidados no OneDrive; verificação rápida de integridade das visualizações; publicação do dashboard.
- Local de publicação: perfil Fiocruz no Tableau Public.
- Motivo da manualidade: licença atual cobre apenas Tableau Desktop e Tableau Prep; não há acesso a Tableau Server/Cloud para automação de publicação.

Governança, segurança e versionamento

- Controle de acesso: permissões no OneDrive restritas a contas autorizadas; acessos ao DW concedidos conforme papéis e responsabilidades.
- Versionamento: salvar cópia diária das extrações consolidadas com timestamp para auditoria e restauração.
- Registro de mudanças: controle de alterações em scripts (.bat), fluxos Power Automate, pipelines no Prep e dashboards, com responsável, data e descrição da modificação.

Resiliência e contingência

- Fallback operacional: procedimento documentado para execução manual do Tableau Prep e reimportação das extrações caso a automação falhe.
- Testes de recuperação: execuções periódicas de testes e simulações de falha para validar procedimentos de restabelecimento.

- Roteiro de comunicação: lista de contatos e responsabilidades para acionamento em incidentes.

Recomendações de evolução

- Avaliar aquisição ou acesso a Tableau Server/Cloud para automação completa de publicação e agendamento de extrações.
- Considerar orquestrador/ETL escalável (Azure Data Factory, Airflow, ou scripts Python) se o volume/complexidade crescer.
- Implementar monitoramento centralizado com painel de status das execuções e métricas de qualidade de dados.
- Automatizar validações e alertas adicionais (ex.: checagens de integridade dos somatórios, thresholds de variação diária).

Siglas e termos

*-1: Dados diários que reflete o fechamento do dia anterior.

Tableau: é uma plataforma de análise visual que ajuda as pessoas e as organizações a obterem o máximo dos seus dados. É uma das ferramentas mais utilizadas para análise de dados, juntamente com o Power BI da Microsoft.

Siglas

Cód.: Código;

COGEAD: Coordenação de Administração;

COGEPLAN: Coordenação de Planejamento;

Desc.: Descrição;

Fiocruz: Fundação Oswaldo Cruz;

GD: Grupo de Despesas;

Núm./Nº: Número;

PI: Plano Interno;

PO: Plano Orçamentário;

PTRES: Programa de Trabalho Resumido;

RAP: Restos a Pagar;

RP: Resultado Primário;

SEOR: Serviços de Orçamentação;

SIAFI: Sistema Integrado de Administração Financeira do Governo Federal;

TED: Termos de Execução Descentralizada;

UG: Unidade Gestora Executora;

UGR: Unidade Gestora Responsável;

UO: Unidade Orçamentária.

Para mais informações ou definições sobre terminologias orçamentárias, consultar o glossário de termos orçamentário (<https://www.congressonacional.leg.br/legislacao-e-publicacoes/glossario-orcamentario>).