



Trabalho Prático 1

Processos ETL em KNIME

Integração de Sistemas de Informação

Licenciatura engenharia de sistemas informáticos (pós-laboral)

Lecionado por: Luís Ferreira

https://github.com/Andrezz1/Projeto_ISI_26012

Autor:

André Reis - 26012



Índice

Introdução	3
Problema	4
Explicação dos Workflows	5
Normalização de Dados	5
Análise das transferências com datas inválidas	6
Análise das transferências pendentes e recusadas	7
Análise das transferências bem sucedidas	8
Criação de Bases de Dados e operações	9
API	10
Vídeo Demonstração (QR CODE)	11
Conclusão	12
Bibliografia	13

Introdução

Este relatório apresenta o desenvolvimento de um trabalho prático no âmbito da disciplina de Integração de Sistemas de Informação, do curso de Engenharia de Sistemas Informáticos. O objetivo principal deste trabalho é aplicar e experimentar técnicas e ferramentas de ETL fundamentais em processos de integração de dados.

Este projeto procura consolidar os conceitos associados à integração de sistemas de informação assim como explorar a framework KNIME para a automação e manipulação eficiente de dados.

Com o crescimento da tecnologia, aumenta a demanda por soluções capazes de integrar dados de diversas fontes e formatos. Neste sentido, os processos ETL assumem um papel essencial.

Ao longo do relatório, serão detalhados os problemas abordados, as estratégias utilizadas, os operadores e as transformações implementadas.

Problema

O trabalho desenvolvido foca-se na resolução de desafios comuns de integração de dados, onde a diversidade de fontes de dados, formatos e necessidades de transformação exige soluções eficientes de ETL.

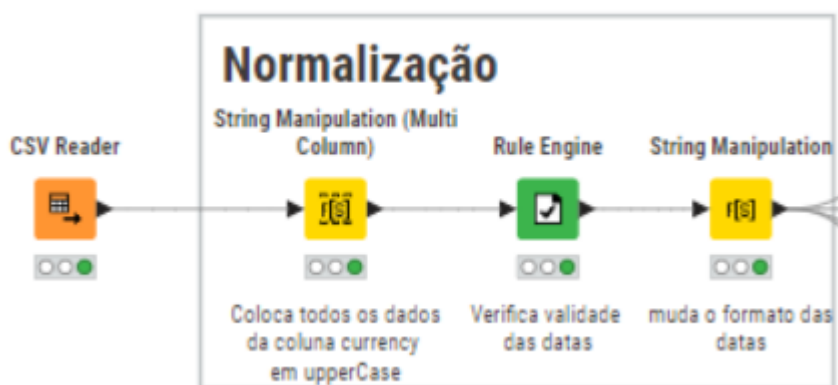
Este projeto visa, portanto, demonstrar a criação e implementação de processos ETL que facilitem a manipulação e integração de dados provenientes de diversas fontes, como base de dados, APIs e ficheiros em formato Json, XML, etc. Pretende-se também resolver problemas comuns como a normalização, limpeza e transformação de dados, de modo a preparar os dados para análises.

Explicação dos Workflows

Normalização de Dados

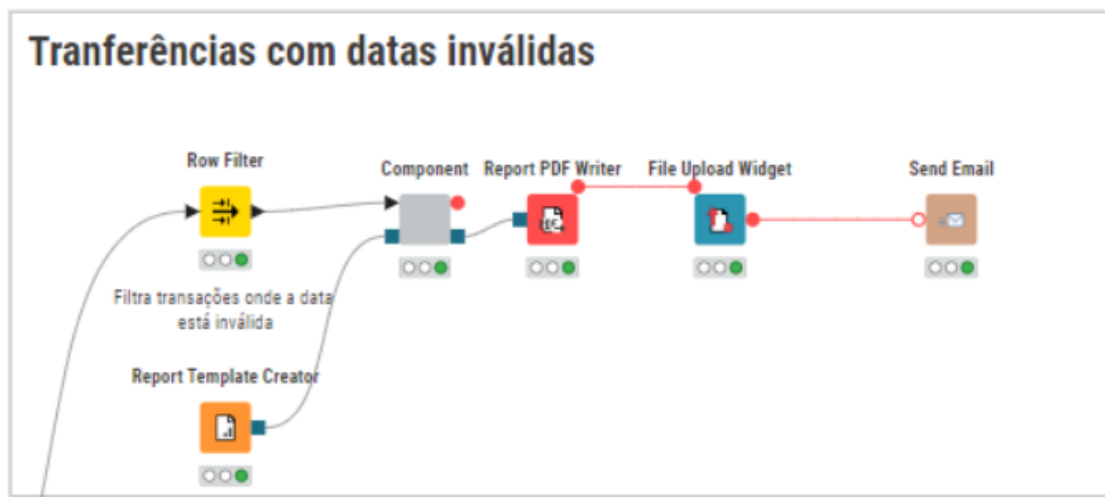
Na imagem abaixo podemos ver um processo de normalização de dados que começa com a leitura de um ficheiro CSV, os dados passam por uma transformação onde os dados da coluna “currency” são convertidos para letras maiúsculas. De seguida, o sistema valida as datas através da utilização de expressões regulares. Finalmente o formato das datas é ajustado para uma forma mais familiar.

Este processo prepara os dados para futuras análises e transformações.



Análise das transferências com datas inválidas

Na imagem abaixo podemos ver uma filtragem das transferências cujas datas foram marcadas como inválidas. De seguida, será gerado um relatório em pdf e o mesmo será enviado por email para correção das datas.



Análise das transferências pendentes e recusadas

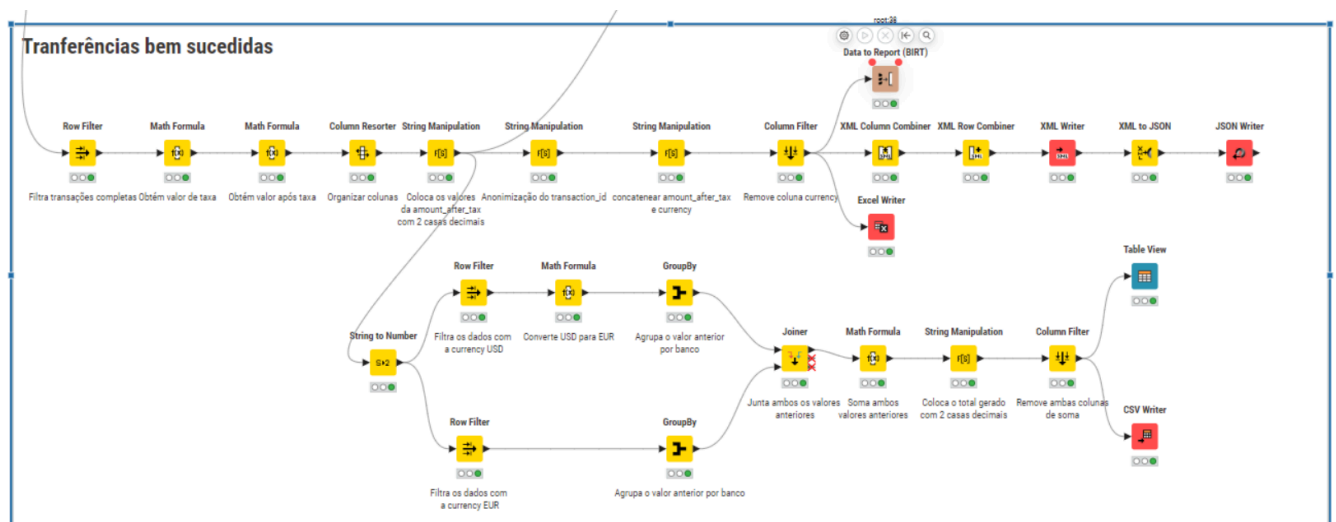
Na imagem abaixo podemos ver um processamento de dados de transferências bancárias recusadas e pendentes. As transações recusadas são agrupadas por banco e um relatório é gerado com esses dados. O relatório final é então exportado em formato pdf, contendo informações sobre as transferências não processadas corretamente.



Análise das transferências bem sucedidas

Na imagem abaixo, são tratadas as transferências bancárias bem sucedidas, são realizadas uma série de operações de transformação e formatação dos dados. Primeiro as transações concluídas são filtradas e ajustadas, aplicando-se cálculos para incluir o valor das taxas e o valor final após a conversão das mesmas. Em seguida, são realizadas operações para organizar as colunas e anonimizar os ids das transações. Após isso os valores são agrupados por banco, e os valores totais em USD(dólar americano) são convertidos para EUR(Euro) para facilitar a comparação.

Os dados processados são exportados para diferentes formatos, Excel, XML, JSON e CSV. O workflow finaliza com a soma dos valores por banco e a preparação do relatório em diferentes formatos, mostrando as transferências bem sucedidas com seus valores totais e outras informações associadas.

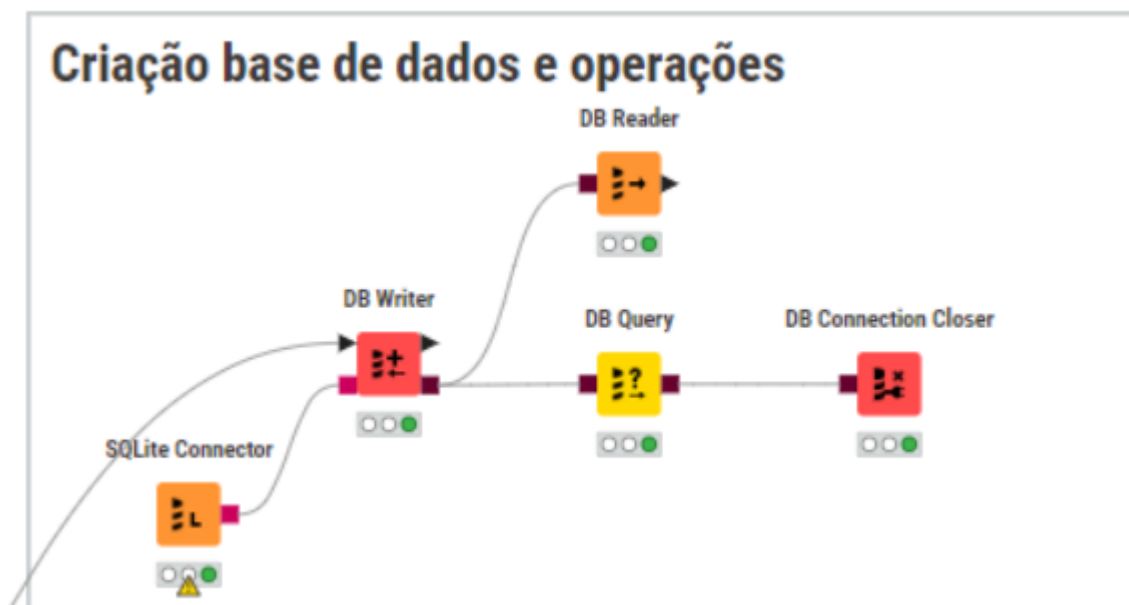


Criação de Bases de Dados e operações

Na imagem abaixo, os dados das transferências são carregados numa base de dados temporária.

Depois de os dados serem inseridos, o “DB Reader” lê as informações da base de dados e mostra uma tabela com todos os dados da base de dados, enquanto o “DB Query” executa uma consulta filtrando apenas as transferências feitas em EUR(Euros). Finalmente o “DB Connection Closer” é usado para fechar a conexão com a base de dados de forma adequada, assim todos os processos são concluídos de forma limpa e segura.

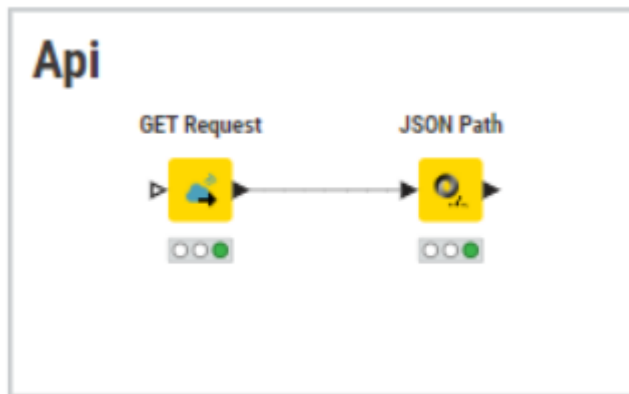
Concluindo este workflow permite o processamento, consulta e visualização das transferências destacando as que foram efetuadas em EUR.



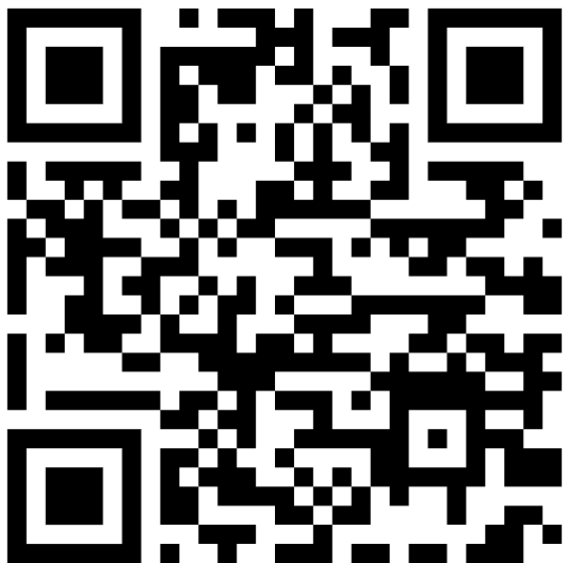
API

Na imagem abaixo, uma requisição é feita a uma API de taxas de conversão de diferentes moedas, que retorna informações sobre o diferente preço das moedas baseado na moeda EUR.

Por fim, os dados JSON são processados, o que permite extrair as informações desse JSON recebido.



Vídeo Demonstração (QR CODE)



Conclusão

Este trabalho prático proporcionou uma oportunidade de explorar e aplicar conceitos de integração de sistemas de informação utilizando a framework KNIME para processos ETL. Através do desenvolvimento das atividades propostas, foi possível enfrentar e resolver desafios comuns na integração de dados.

A implementação de processos ETL eficientes demonstrou-se fundamental para lidar com problemas de normalização, limpeza e transformação de dados, o que resultou num fluxo de dados otimizado e mais confiável. Adicionalmente, o uso de KNIME permitiu automatizar diversas etapas do processo, facilitando a gestão e integração de dados de diferentes formatos, como JSON, XML.

Bibliografia

ChatGPT: <https://chat.openai.com/>

Api das taxas de conversão das moedas: <https://exchangeratesapi.io/>

Gerador de QR code: <https://online-qr-generator.com/>

Documentação KNIME: <https://docs.knime.com/>