Trabalho prático 1

Integração de Sistemas de Informação



José Moura – a13742 Paulo Meneses – a17611

Introdução

Este trabalho foi realizado no âmbito da unidade curricular, Integração de Sistemas de Informação, no curso de Licenciatura em Engenharia em sistemas informáticos pós-laboral.

No trabalho foi desenvolvido soluções de ETL (Extract, Transform, Load) com a ferramenta Pentaho Kettle, que permite a criação de diversos processos de ETL, neste caso o envio de um email com a previsão meteorológica de braga.

Objetivos no âmbito da disciplina

• Consolidar conceitos associados à Integração de Sistemas de Informação usando Dados;

• Analisar e especificar cenários de aplicação de processos de ETL;

• Explorar ferramentas de suporte a processos de ETL;

• Explorar novas Tecnologias, Frameworks ou Paradigmas;

• Potenciar a experiência no desenvolvimento de software;

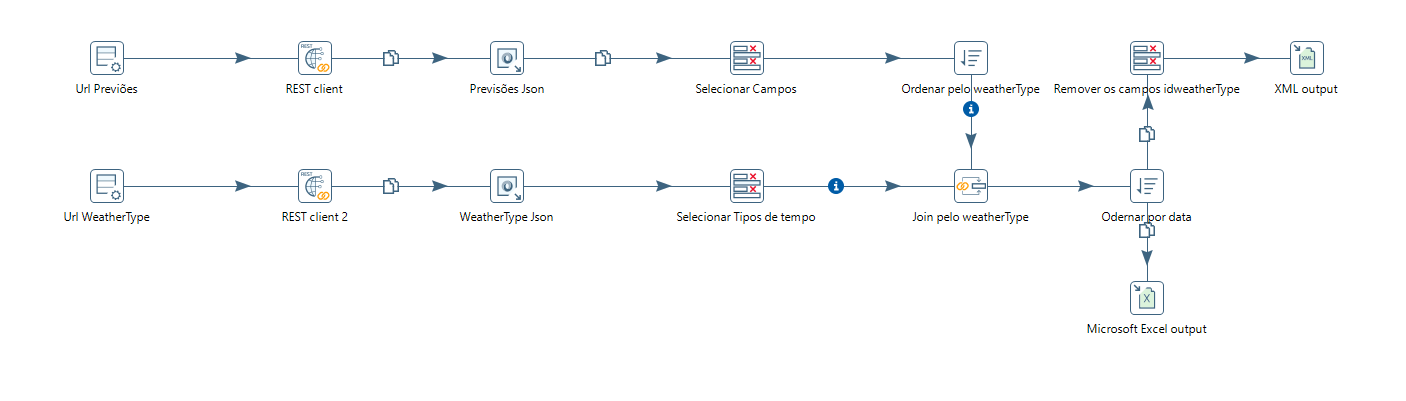
• Facilitar a assimilação do conteúdo da Unidade Curricular.

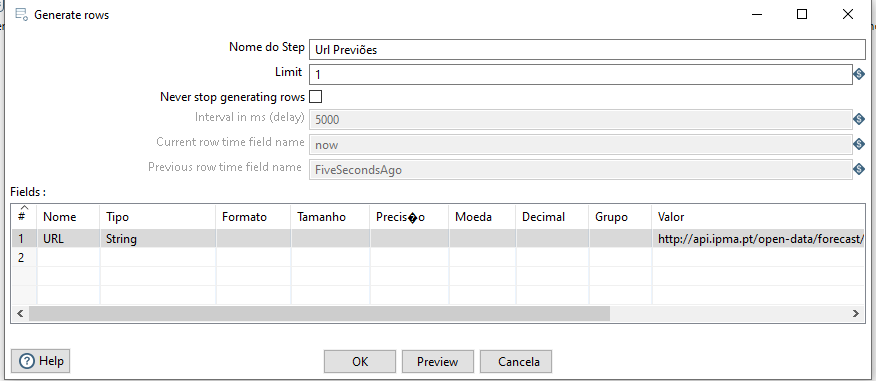
Resumo

Este trabalho foi desenvolvido na ferramenta Pentaho Kettle, na qual o resultado do processo ETL é o envio de um email com a previsão meteorológica de Braga para os próximos 5 dias.

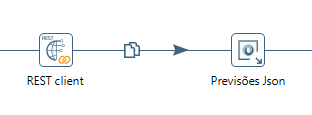
Os dados utilizados são retirados da API do IPMA (Instituto Português do Mar e da Atmosfera).

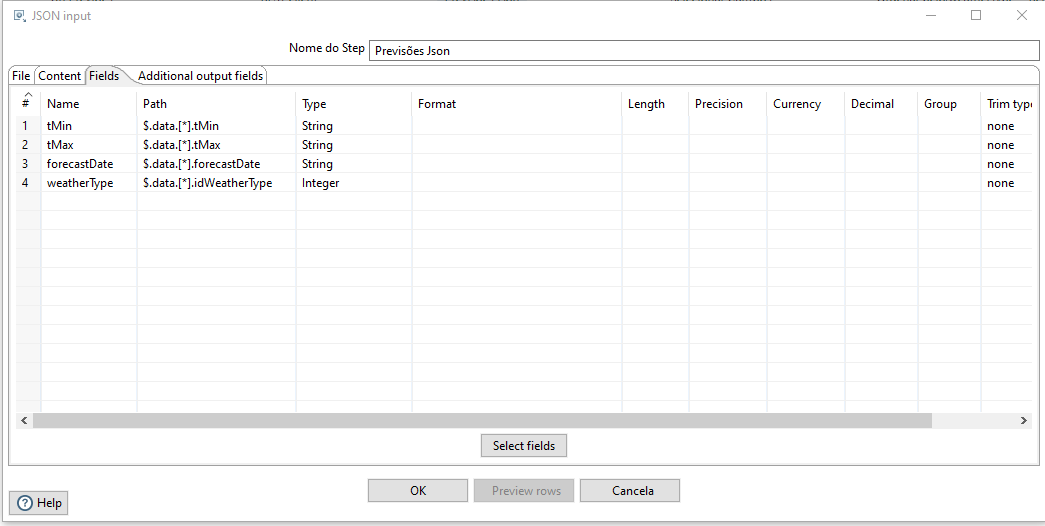
Desenvolvimento

1. Transformações
   1. Primeira Transformação

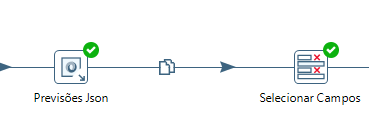
É utilizado um Generate Rows (Url Previsões) para gerar um linha com o link (<http://api.ipma.pt/open-data/forecast/meteorology/cities/daily/1030300.json>) para as previsões de braga.

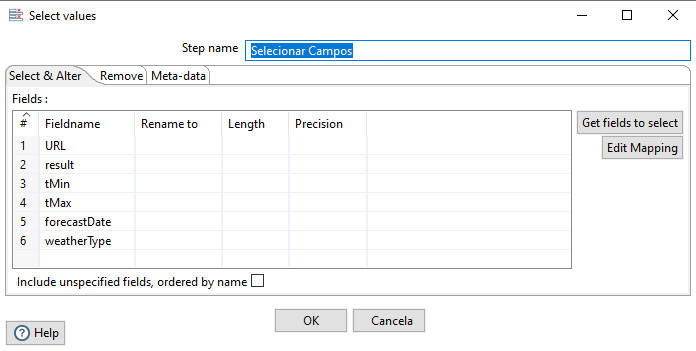
Depois o Rest Client utiliza esse link para receber o json através do metodo GET e o output é inserido num Json Input (Previsões Json).

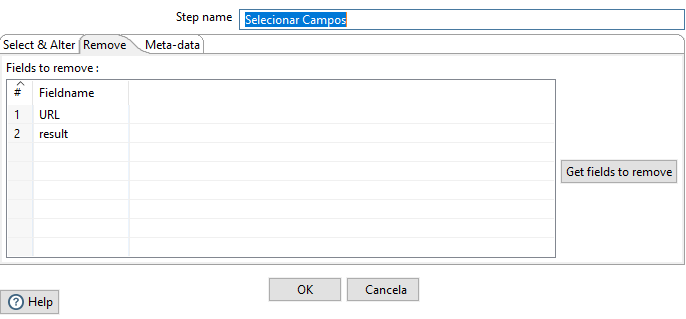


Os campos do Previsões Json que foram configurados foram: tMin, tMax, forecastDate e weatherType.

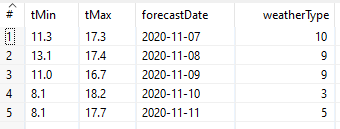
O output desse json é inserido num Select Values (Selecionar Campos) para ser selecionado somente os valores de tMin, tMax, forecastDate e weatherType.



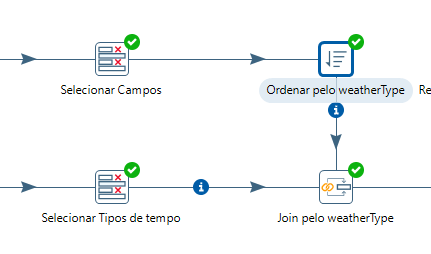




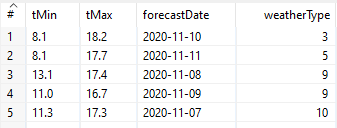
O resultado deste Select Values é o seguinte:

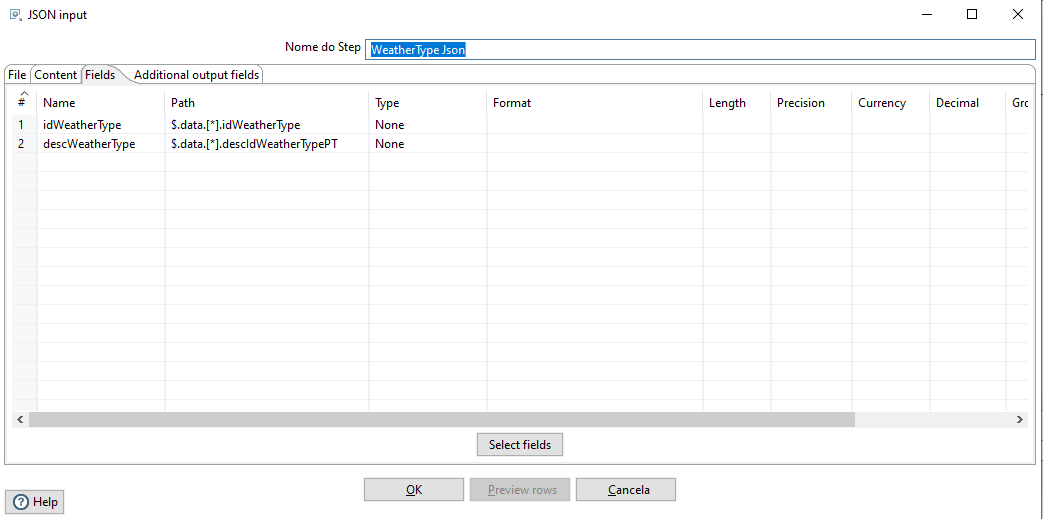


Os valores são ordenados pela ordem crescente pelo campo weatherType no Sort Rows (Ordenar pelo weatherType) para depois serem juntados a uma tabela com a descrição para cada tipo de tempo.

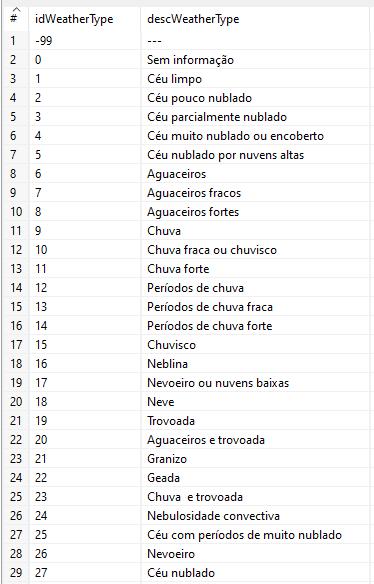
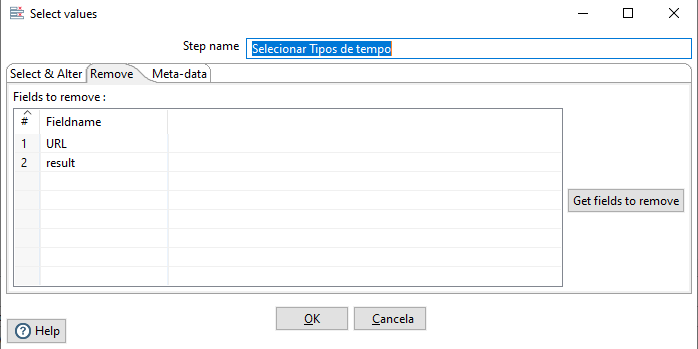


Depois da ordenação a tabela resultado é a seguinte:

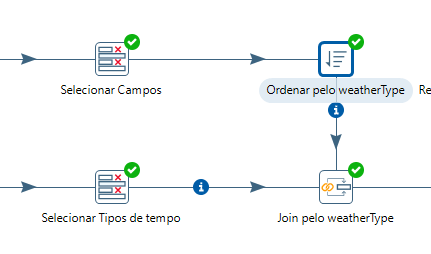


O processo de obter a descrição dos tipos de tempo é semelhante ao anterior, é criado uma linha com um Generate Rows (Url weatherType), com o link (<http://api.ipma.pt/open-data/weather-type-classe.json>), depois o Rest Client 2 efetua o download do json, os campos inseridos no Json Input (WeatherType Json) são: idWeatherType e desWeatherType.

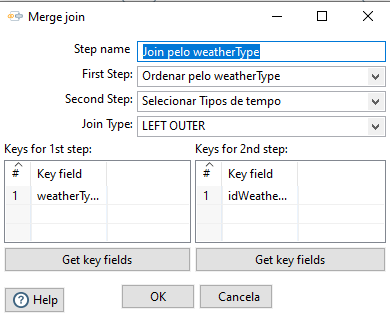
Depois é utilizado um Select Values (Selecionar Tipos Tempo) para excluir os campos URL e result.

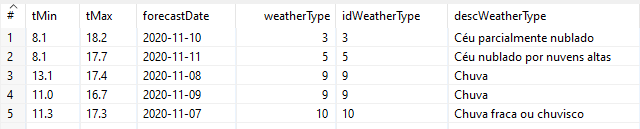
O tabela resultado dos tipo de tempo é:

No fim o resultado é inserido no Merge Rows (Join pelo weatherType) visto que ele já está ordenado.

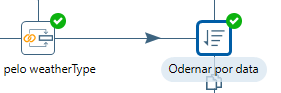


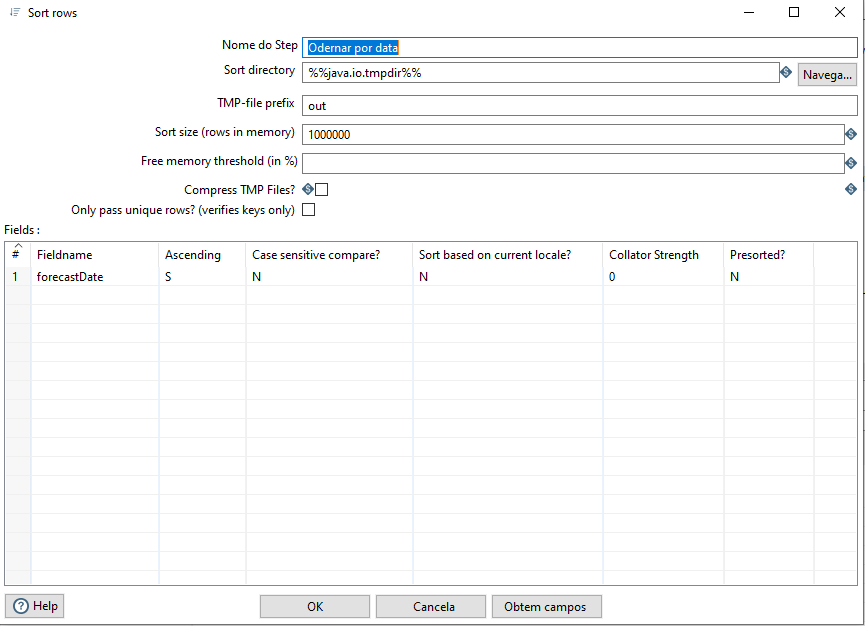
O Merge Rows permite fazer o cruzamento dos dados de uma tabela com a outra por determinado valor chave, neste caso o weatherType.



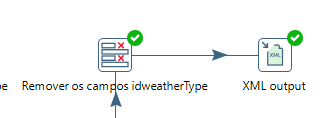
O resultado do Merge Rows é:

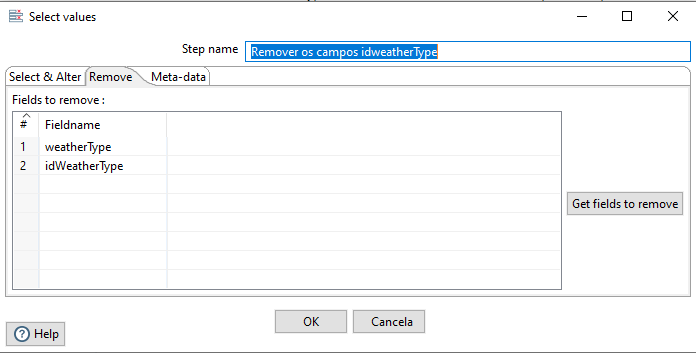
Esta tabela é de seguida ordenada por ordem crescente pelo campo forecastDate, no Sort Rows (Ordenar por data).

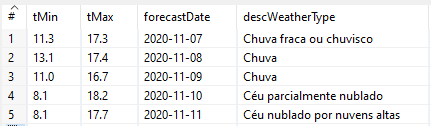




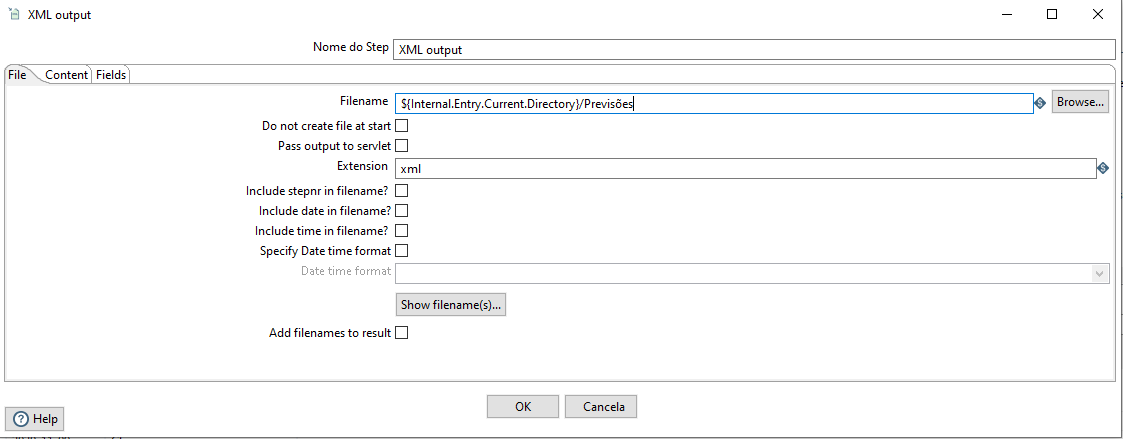
Depois é efetuada a remoção dos campos weatherType e idWeatherType através do Select Values (Remover os campos idweatherType) e o output é enviado para um ficherio xml.

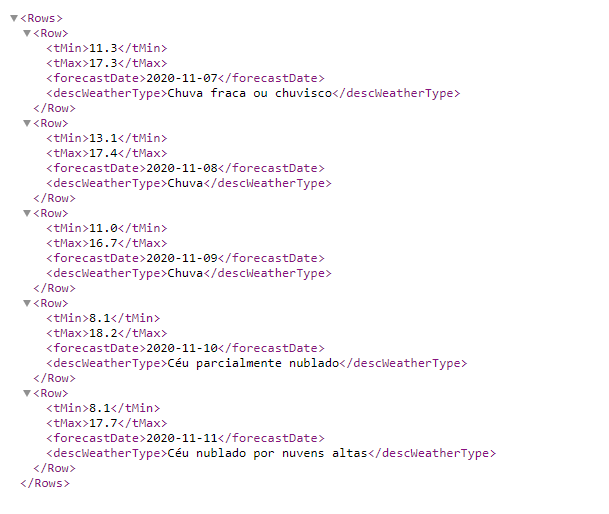


Remoção dos campos weatherType e IdWeatherType

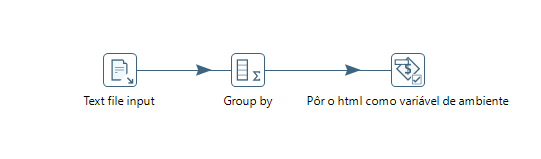
A tabela que resulta é :

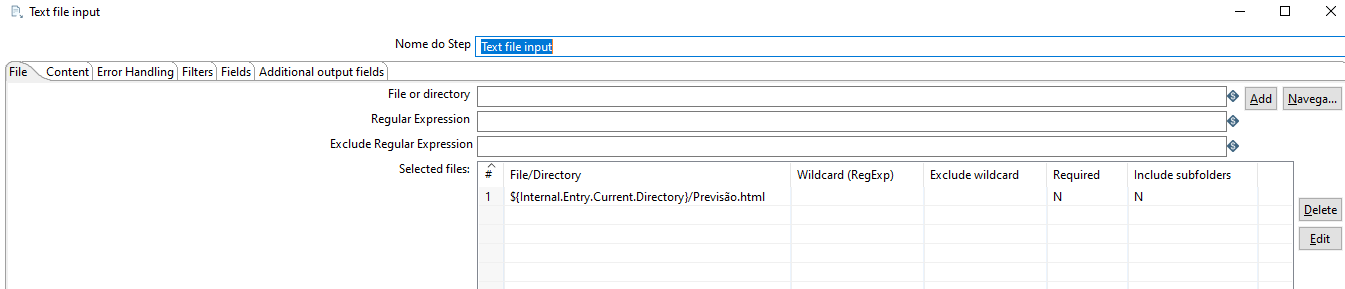
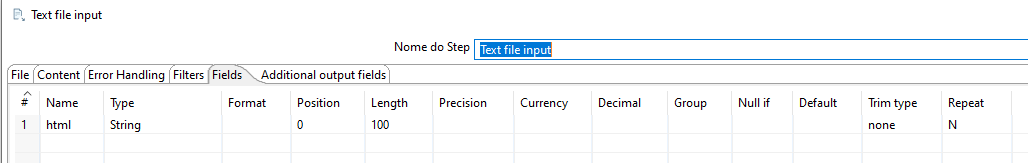
No fim é gerado um xml com o nome de Previsões que contem a tabela.

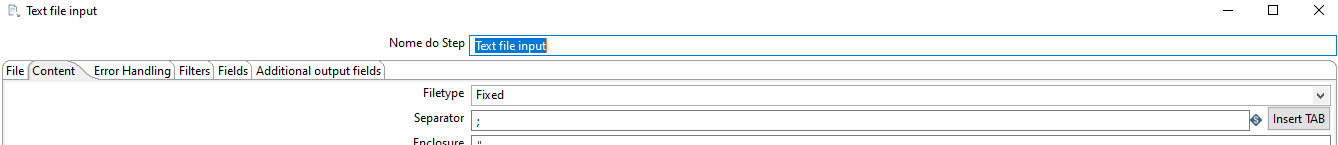


O conteúdo do xml é o seguinte:

1.2 Tranformação 2

 A segunda transformação é pegar no conteúdo de um ficheiro html e defini-lo como variável de ambiente.

 O ficheiro é lido como um ficheiro de texto, neste caso o ficheiro lido é Previsão.html, na configuração do Content é preciso mudar o fileType para Fixed e nos Fields é preciso a configuração do campo com Name html, Type string, Position 0, Length 100, Trim Type None e Repeat N.



É utilizado um Group by para a concatenação das strings lidas numa só e depois essa string final é passada para um Set Variables (SetStringAmbient) para pôr a string como variável de ambiente para ser passada para o email.

