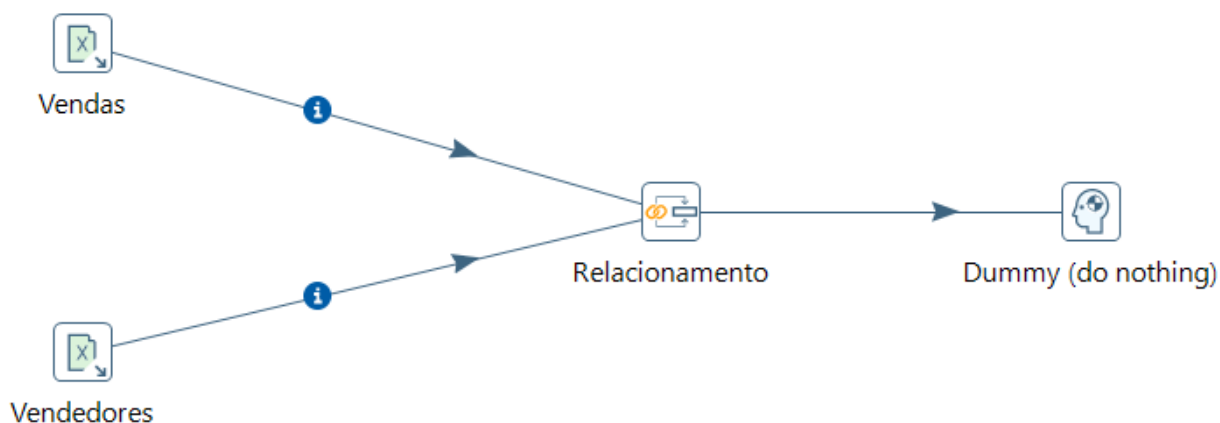
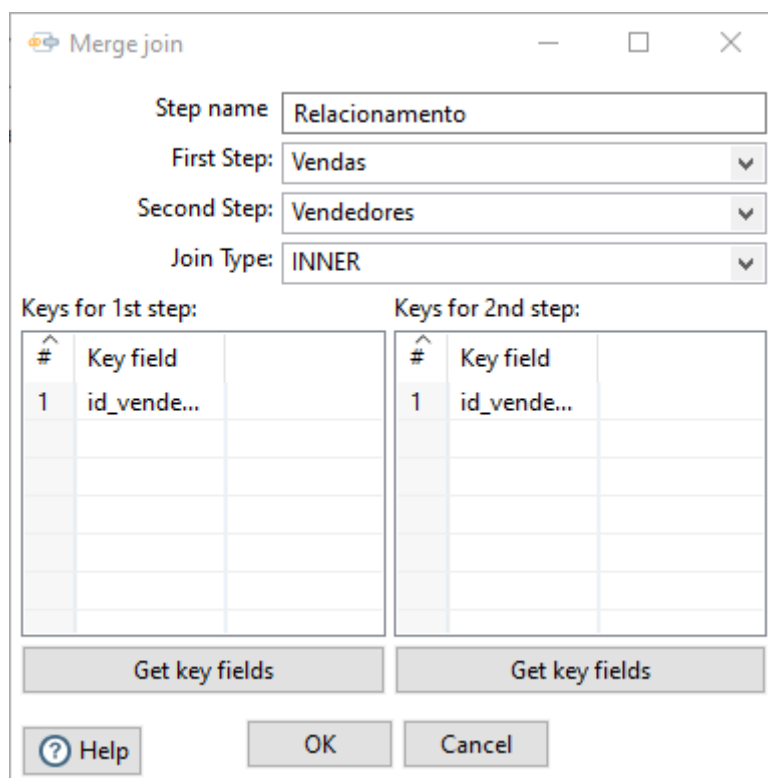


Exercício 01 – Step Merge Join

- 1.1** Crie uma pasta chamada LAB03 em seu repositório, crie uma nova transformação utilizando os inputs “vendas.xlsx”, “vendedores.xlsx”, crie um step do tipo “Merge Join”, faça o vínculo entre os dois arquivos e execute:



- 1.2** Exemplo de configuração do step “Merge Join”:



A janela de configuração do step 'Merge join' mostra as seguintes configurações:

- Step name: Relacionamento
- First Step: Vendas
- Second Step: Vendedores
- Join Type: INNER

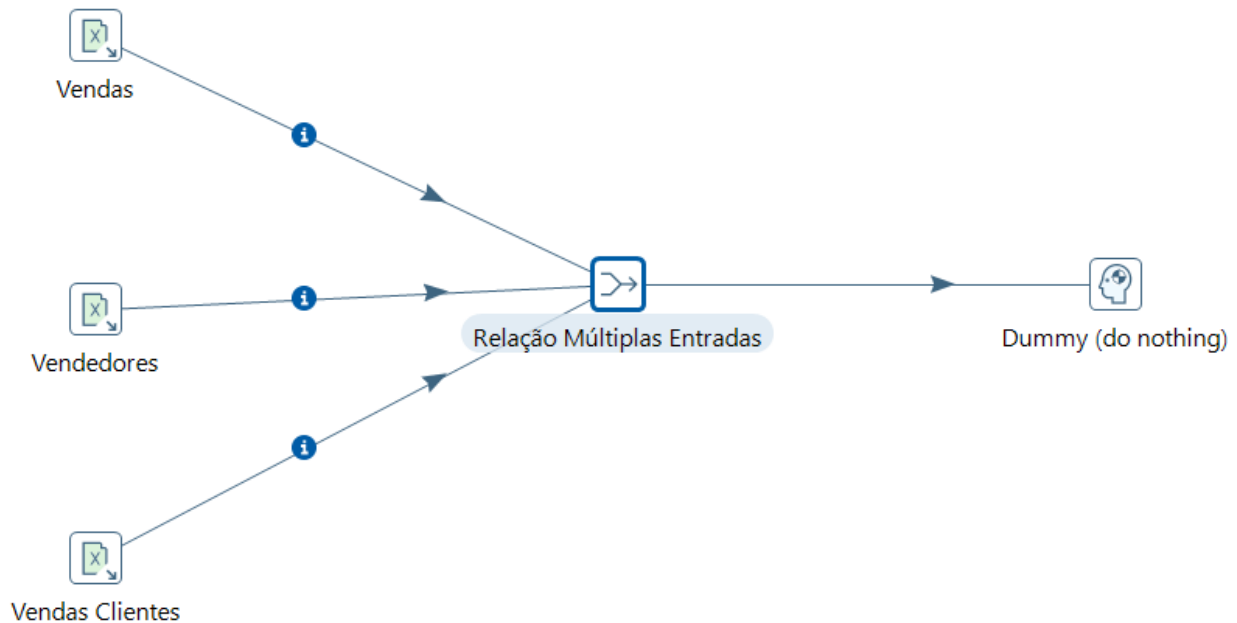
Abas para configuração de chaves:

Keys for 1st step:		Keys for 2nd step:	
#	Key field	#	Key field
1	id_vende...	1	id_vende...

Botões: Get key fields, Get key fields, Help, OK, Cancel.

Exercício 02 – Step Multiway Merge Join

2.1 Crie uma transformação nova, adicione os inputs “vendas.xlsx”, “vendedores.xlsx” e “vendas_clientes.xlsx”. Em seguida, adicione o step “Multiway Merge Join”, faça a relação entre todos os inputs, gere a saída em um dummy. Salve a transformação:



2.2 Exemplo de configuração do step “Multiway Merge Join”:

Multiway merge join

Step name	Relação Múltiplas Entradas		
Input Step1	Vendas	Join Keys	id_vendedor
Input Step2	Vendedores	Join Keys	id_vendedor
Input Step3	Vendas Clientes	Join Keys	id_vendedor
	INNER		

Help

OK

Cancel

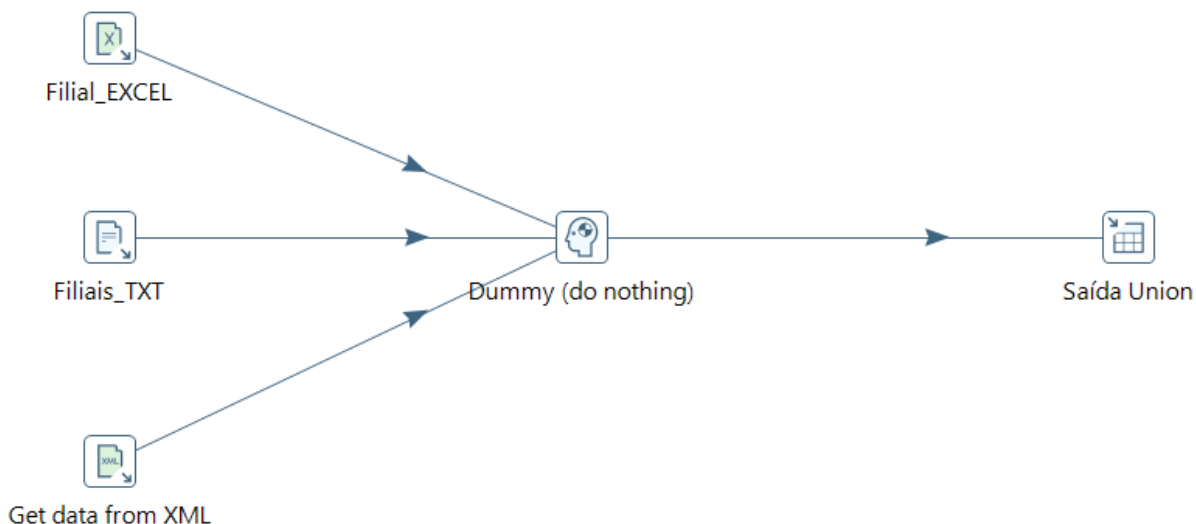
Select Keys

Select Keys

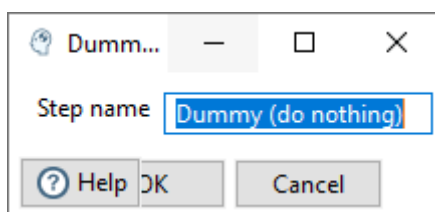
Select Keys

Exercício 03 – Union All Dummy

3.1 Crie uma transformação nova, adicione os inputs “vendas.xlsx”, “vendedores.xlsx” e “filial_01.xls”, “filial_02.txt” e “filial_03.xml”. Em seguida, adicione o step “Dummy”, faça a união entre todos os arquivos de entrada e grave seu output no Postgres, em uma tabela chamada “saída_union”. Salve e execute a transformação:



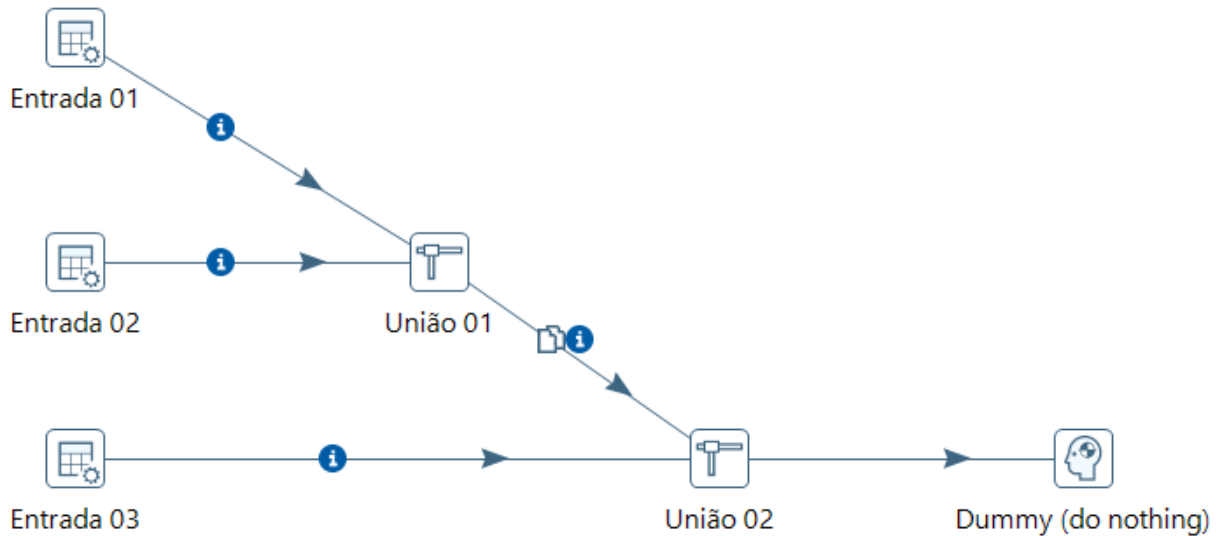
3.2 Exemplo de configuração do step “Dummy”:



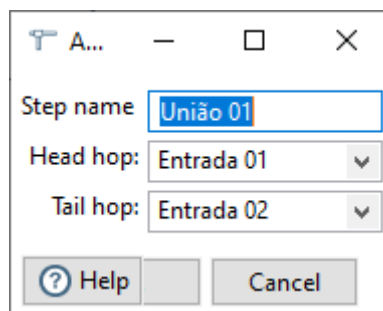
Não há o que se configurar no Dummy, apenas o nome

Exercício 04 – Append Streams

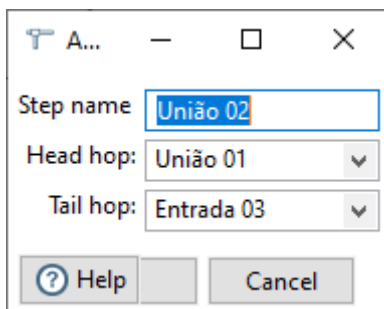
4.1 Crie uma transformação nova, adicionando 3 inputs do tipo “Data Grid” e os preencha com dados aleatórios. Em seguida, adicione o step “Append Streams” e faça a junção dos dois primeiros steps. A saída deste step de junção, será adicionada ao segundo step de junção, em conjunto com a terceira entrada “Data Grid”. Gere a saída para um Dummy e salve a transformação:



4.2 Exemplo de configuração do step “Append Streams (União 01)”:



4.3 Exemplo de configuração do step “Append Streams (União 02)”:



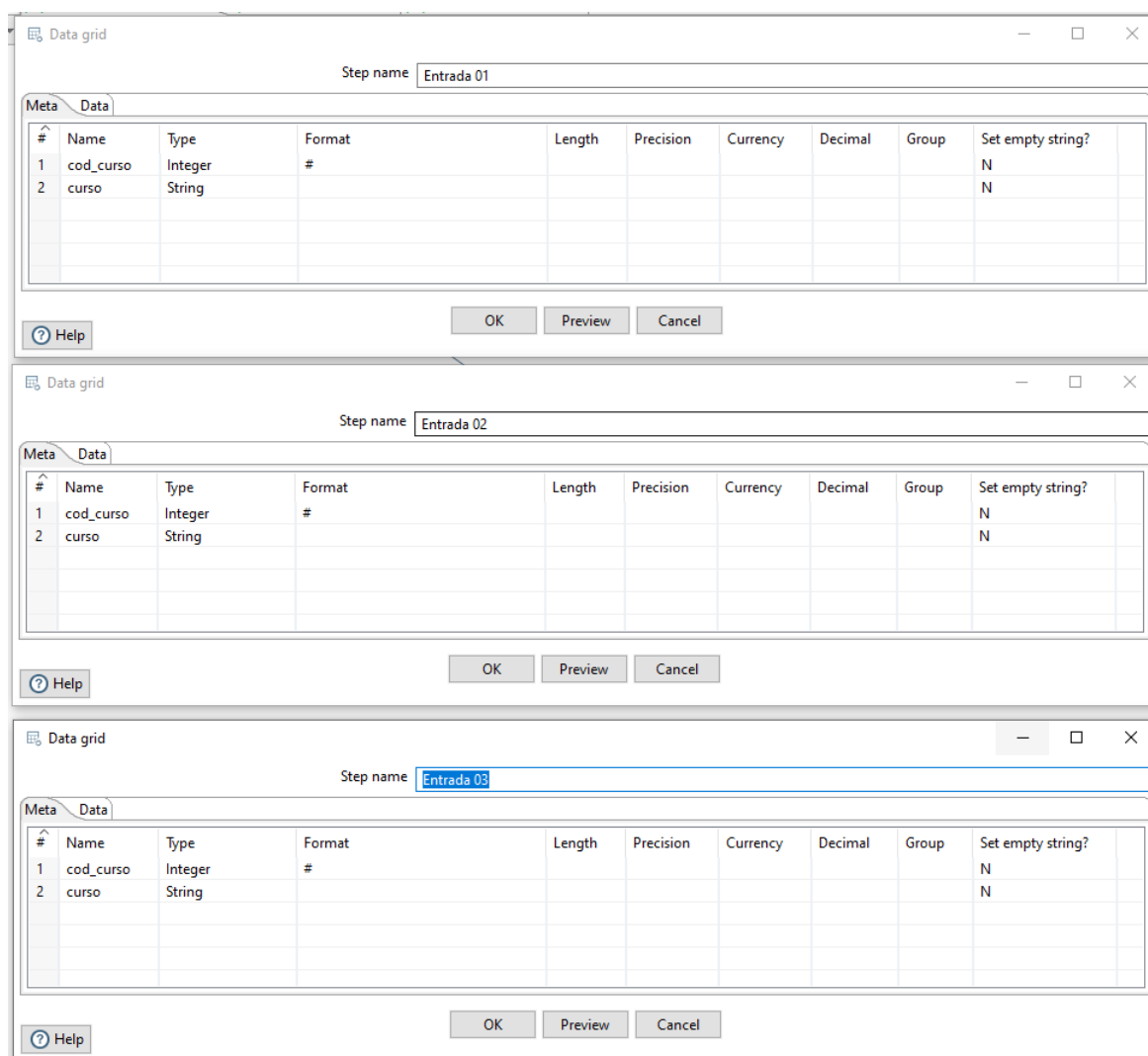
Step name: União 02

Head hop: União 01

Tail hop: Entrada 03

Buttons: Help, Cancel

4.4 Os steps de entrada devem possuir a mesma estrutura:



Three screenshots of the 'Data grid' window for steps 'Entrada 01', 'Entrada 02', and 'Entrada 03'. Each window shows a table with columns: #, Name, Type, Format, Length, Precision, Currency, Decimal, Group, and Set empty string?.

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Set empty string?
1	cod_curso	Integer	#						N
2	curso	String							N

4.5 Sugestão de dados para os inputs “Data Grid”

Data grid

Step name

Meta

Data

#	cod_curso	curso
1	101	Direito
2	201	Administração
3	301	Engenharia
4	401	Odontologia

Help

OK

Preview

Cancel

Data grid

Step name

Meta

Data

#	cod_curso	curso
1	501	Economia
2	601	Nutrição
3	301	Engenharia
4	401	Odontologia

Help

OK

Preview

Cancel

Data grid

Step name

Meta

Data

#	cod_curso	curso
1	701	Hotelaria
2	801	Medicina
3	901	Odontologia
4	401	Odontologia

Help

OK

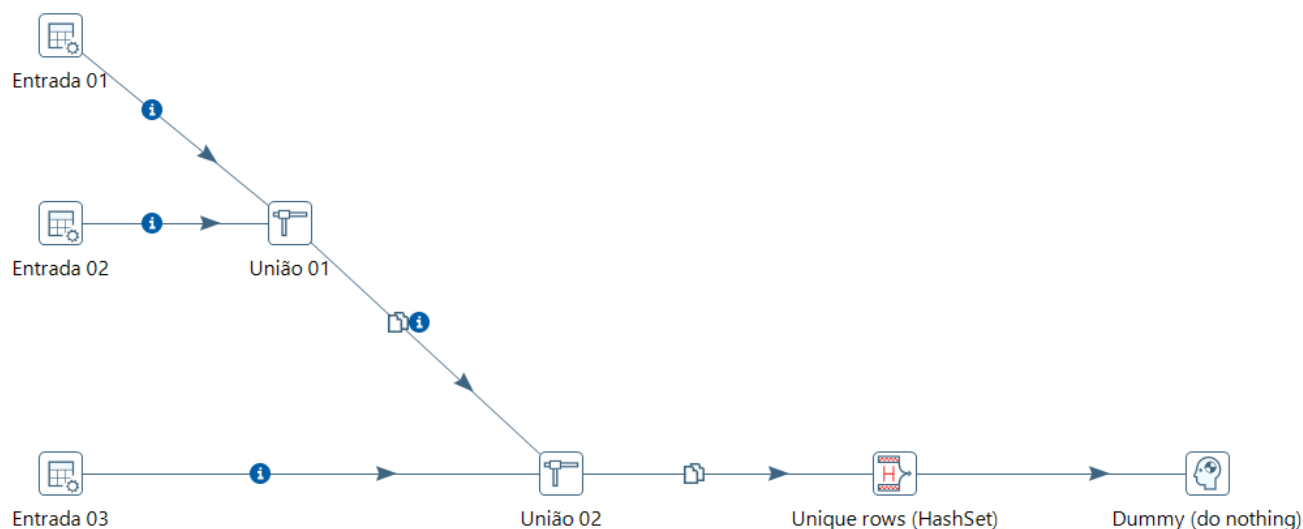
Preview

Cancel

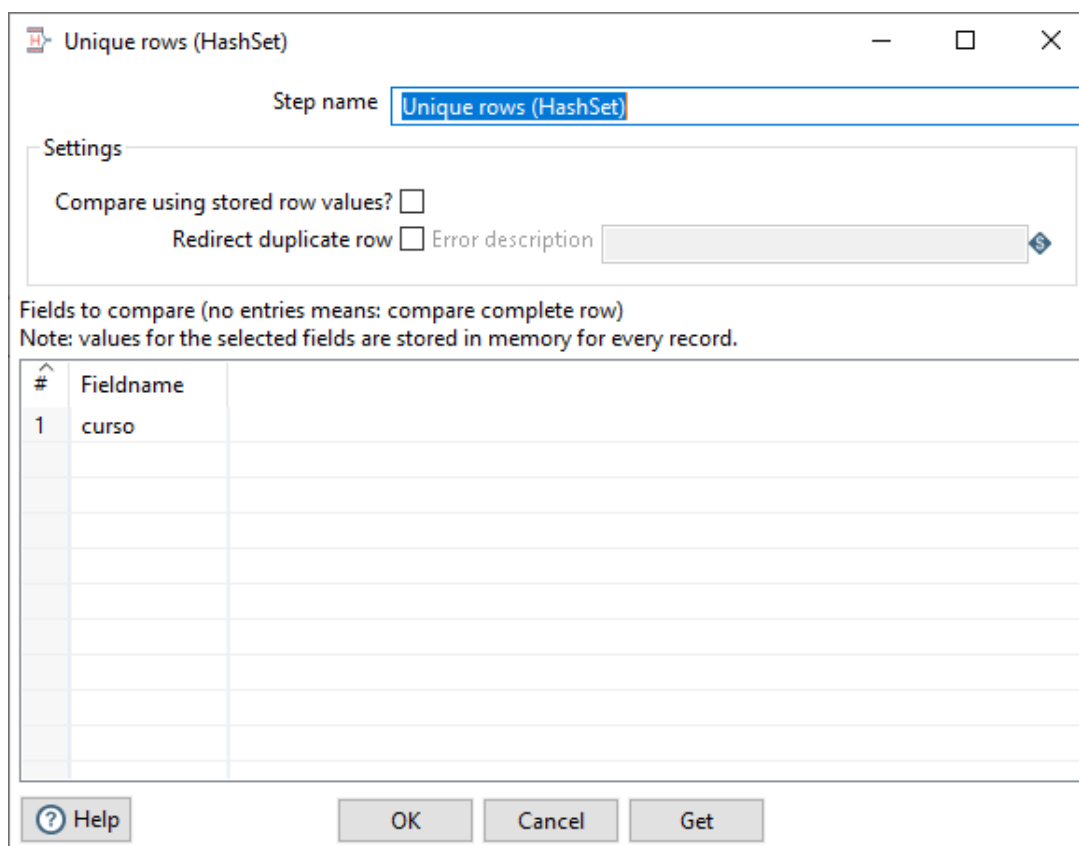
4.6 Salve a transformação e execute:

Exercício 05 – Unique Rows

5.1 Repita os passos do exercício 4, adicionando ao final o step “Unique rows (HashSet)”.
Salve como uma nova transformação e execute.



5.2 Exemplo de configuração do step “Unique rows(HashSet)”:

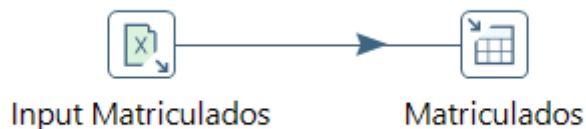


The screenshot shows the configuration window for the 'Unique rows (HashSet)' step. The 'Step name' field is set to 'Unique rows (HashSet)'. Under the 'Settings' section, the 'Compare using stored row values?' checkbox is unchecked, and the 'Redirect duplicate row' checkbox is also unchecked. The 'Error description' field is empty. Below the settings, there is a section for 'Fields to compare (no entries means: compare complete row)' with a note: 'Note: values for the selected fields are stored in memory for every record.' A table with two columns, '#', and 'Fieldname', contains one entry: '1' and 'curso'. At the bottom of the window, there are buttons for 'Help', 'OK', 'Cancel', and 'Get'.

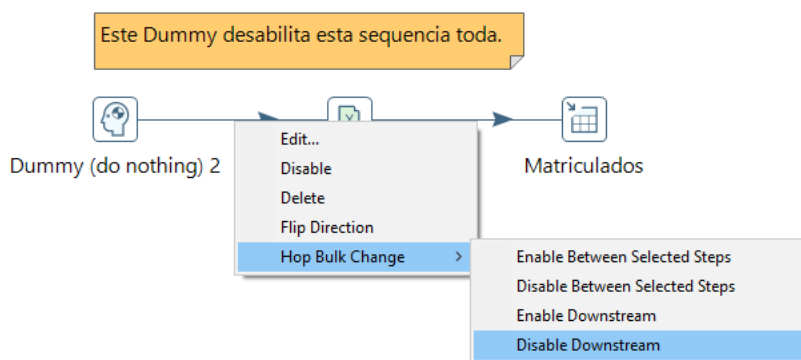
#	Fieldname
1	curso

Exercício 06 – Merge rows (diff)

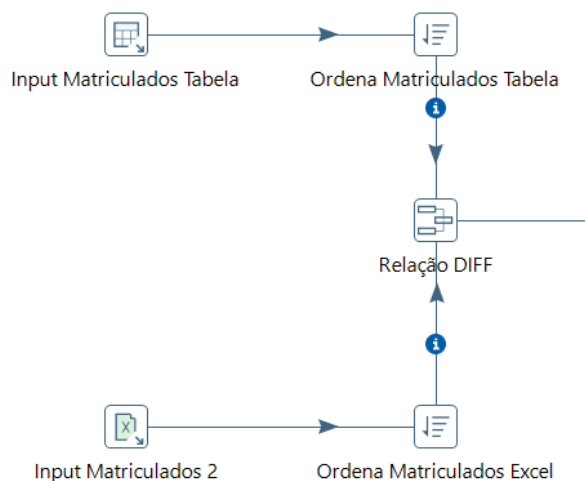
6.1 Crie uma nova transformação, e a salve como “06 – Busca Merge rows (diff)”. Em seguida, os dados de “matriculados.xlsx”, e o salve no Postegres em uma tabela chamada “Matriculados”.



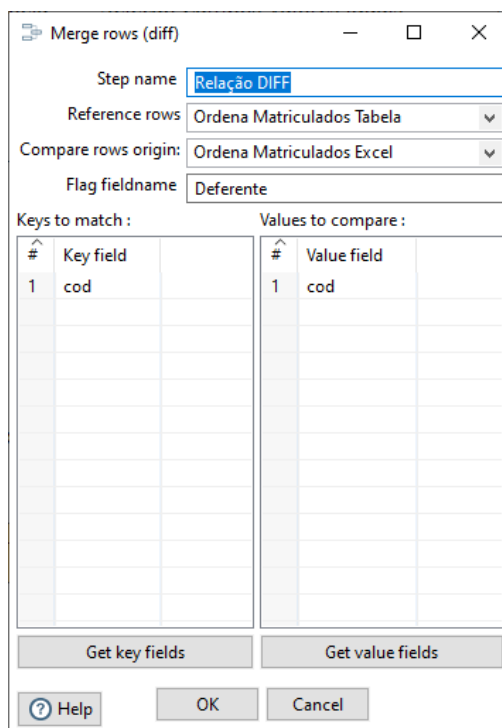
6.2 Após a execução, com sucesso, adicione um “Dummy” ao início e desabilite os HOPS, para desabilitar todo o fluxo.



6.3 Na mesma transformação, crie um novo fluxo, lendo a tabela “Matriculados” recém carregada e também lendo o arquivo “matriculados.xlsx” (não utilizar o step já criado, criar um novo. Após isto, ordenar ambos os inputs pelo campo “cod” em ordem crescente. Ligar as duas saídas já ordenadas ao step “Merge rows (diff)”:



6.4 Exemplo de configuração do step “Merge rows (diff)”. Observe que o campo “Flag fieldname” será criado, e irá indicar se o registro é novo ou idêntico:



Step name:

Reference rows:

Compare rows origin:

Flag fieldname:

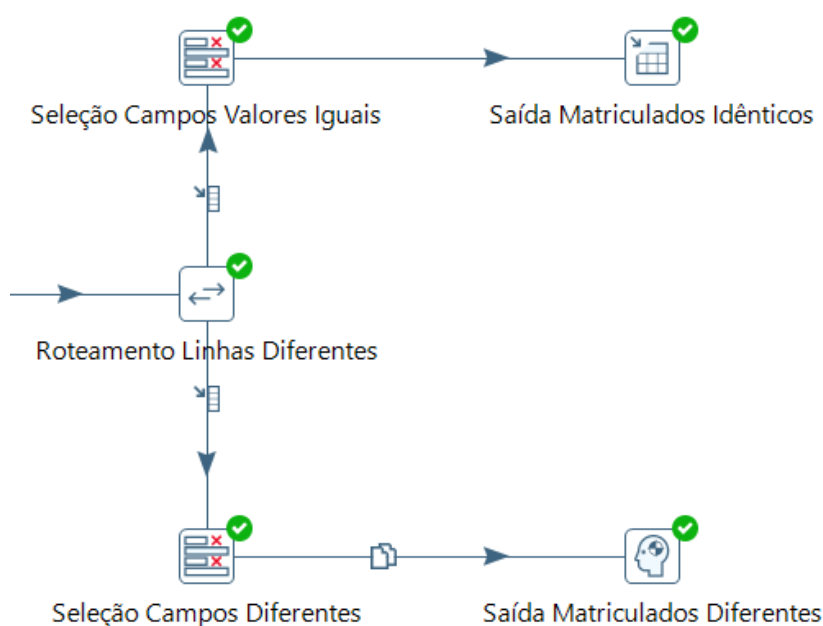
#	Key field
1	cod

#	Value field
1	cod

Get key fields Get value fields

Help OK Cancel

6.5 Adicione um step “Switch Case” para separar os registros novos dos existentes, e grave os campos iguais em uma nova tabela chamada “matriculados_identicos”, levando os campos diferentes para o step “Dummy”:



6.6 Salve o resultado final e edite o arquivo “matriculados.xlsx” para nova execução:

11	32163	IFB	Educação Presencial	TÉCNICO	Subsequente	2009	matriculado	291
12	32164	CBRA	Educação Presencial	ESPECIALIZACAO (LATO SENSU)	Outros	2010	matriculado	112
13	32165	CBRA	Educação Presencial	FORMACAO CONTINUADA	Outros	2010	matriculado	366
14								
15								
16								

Arquivo original

10	32162	IFB	Educação Presencial	TÉCNICO	Integrado	2009	matriculado	182
11	32163	IFB	Educação Presencial	TÉCNICO	Subsequente	2009	matriculado	291
12	32164	CBRA	Educação Presencial	ESPECIALIZACAO (LATO SENSU)	Outros	2010	matriculado	112
13	32165	CBRA	Educação Presencial	FORMACAO CONTINUADA	Outros	2010	matriculado	366
14	32166	UFU	Educação Presencial	ESPECIALIZACAO (LATO SENSU)	Outros	2010	matriculado	400
15	32167	UFU	Educação Presencial	FORMACAO CONTINUADA	Outros	2010	matriculado	500
16								

Arquivo com adição de duas novas linhas