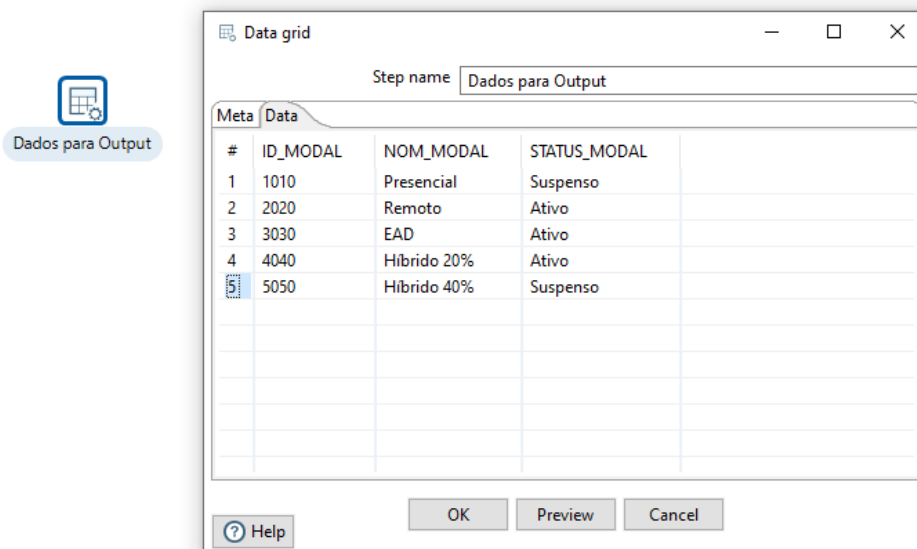
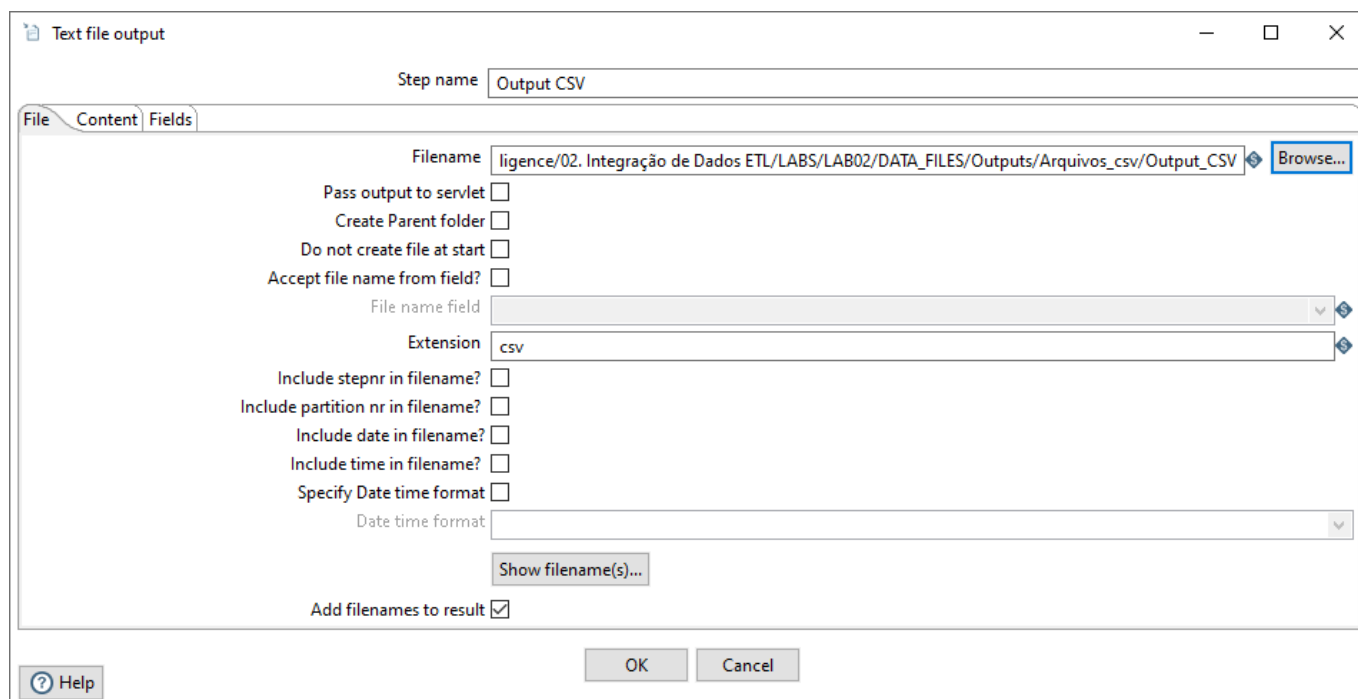


Exercício 08 – Output CSV

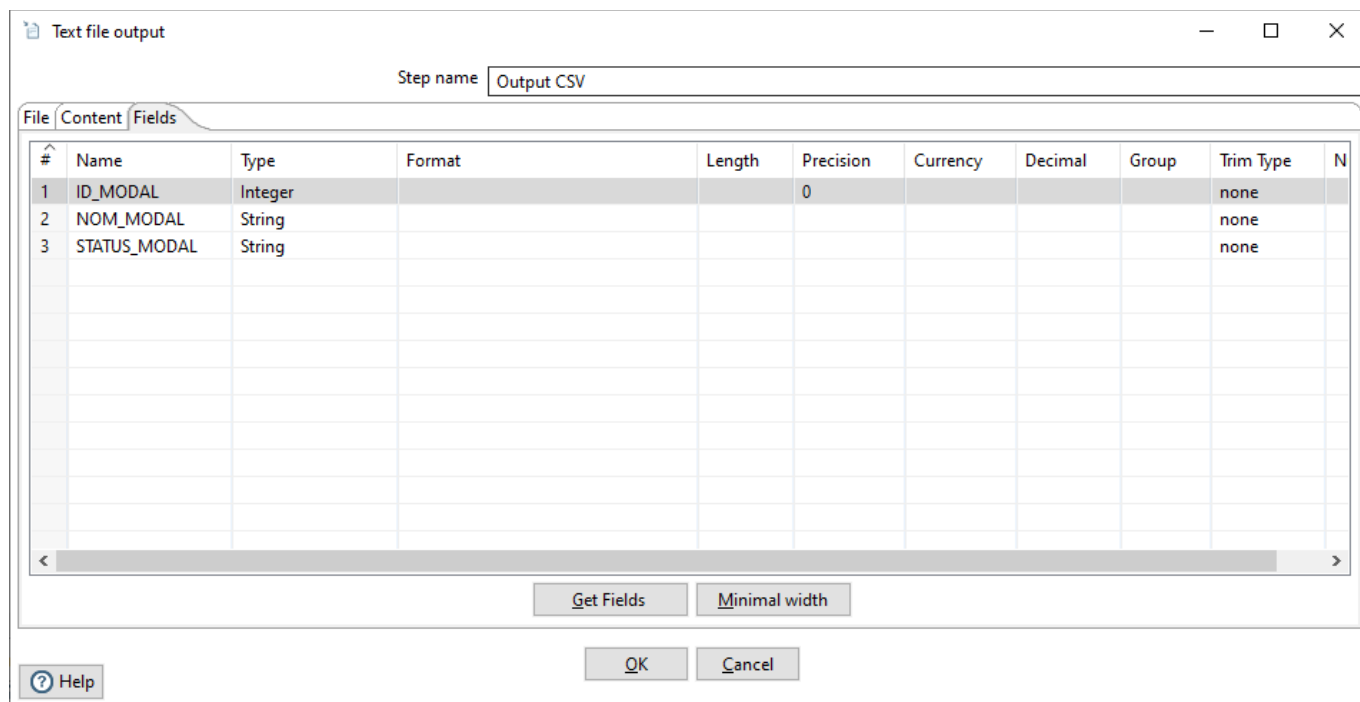
8.1 Crie uma nova transformação e a salve no repositório como “08 – Output CSV”. Crie um step “Data Grid” para servir como a entrada de dados:



8.2 Configure o step de output; na aba “file”, indique o caminho de destino e nome do arquivo em “Filename”, insira “csv” no campo “Extension”:

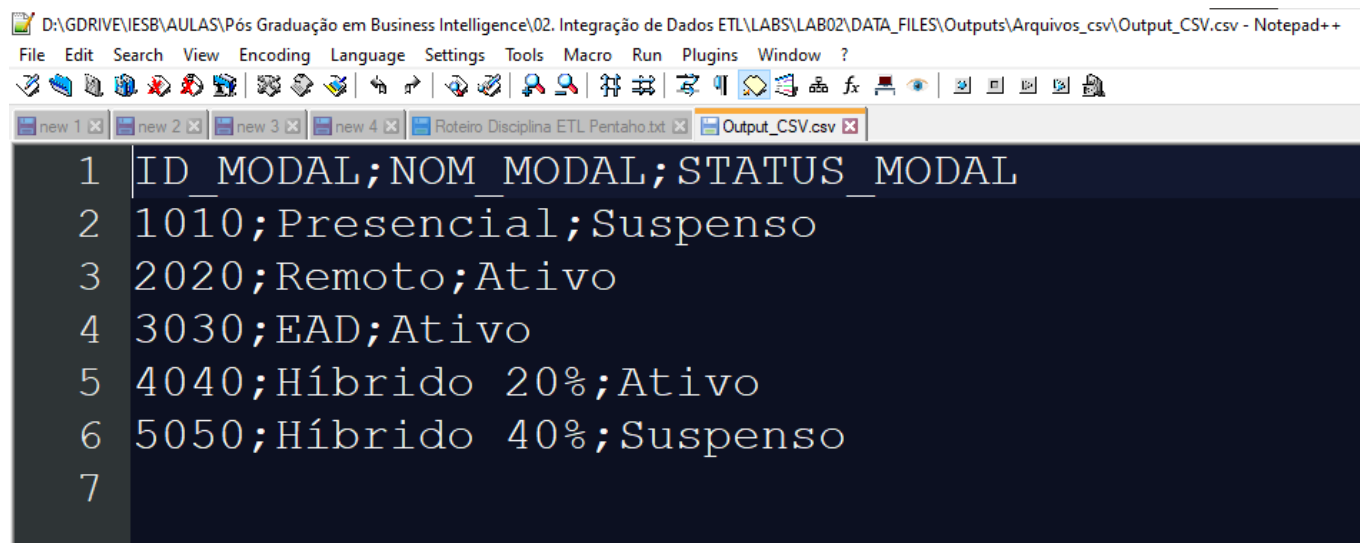


8.3 Na aba “Fields”, clique em “Get fields” para criar a estrutura necessária à criação do arquivo na pasta indicada:



#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim Type	N
1	ID_MODAL	Integer			0				none	
2	NOM_MODAL	String							none	
3	STATUS_MODAL	String							none	

8.4 Execute a transformação, navegue até a pasta indicada para criação do arquivo e valide se sua criação foi concluída com sucesso:



```

1 ID_MODAL;NOM_MODAL;STATUS_MODAL
2 1010;Presencial;Suspendo
3 2020;Remoto;Ativo
4 3030;EAD;Ativo
5 4040;Híbrido 20%;Ativo
6 5050;Híbrido 40%;Suspendo
7

```

Exercício 09 – Output TXT

9.1 Repita as etapas do Exercício 7, realizando as alterações necessárias à criação de um arquivo com extensão TXT:

DICA: Mantendo a transformação “08 – Output CSV” aberta, crie uma nova transformação com o nome “09 – Output TXT”. Volte à transformação do arquivo csv, clique com o botão direito no step “Dados para Output”, copie este step e o cole na área gráfica da transformação TXT. Este procedimento pode ser realizado em diversas situações para reaproveitar steps entre transformações.

Exercício 10 – Output XML

10.1 Repita as etapas do Exercício 8, realizando as alterações necessárias à criação de um arquivo com extensão XML.

Exercício 11 – Output Table

11.1 Crie uma nova transformação, e a salve como “11 – Output Table”. Insira o step anterior “Dados para Output” e o step “Table Output”, criando o HOP entre estes dois steps. Não é necessário criar a tabela de destino em banco de dados, pois ela será criada dentro do step de saída da tabela relacional. Configure inicialmente o “Step name”, “Connection”, “Target schema” e o nome da tabela que se deseja criar:

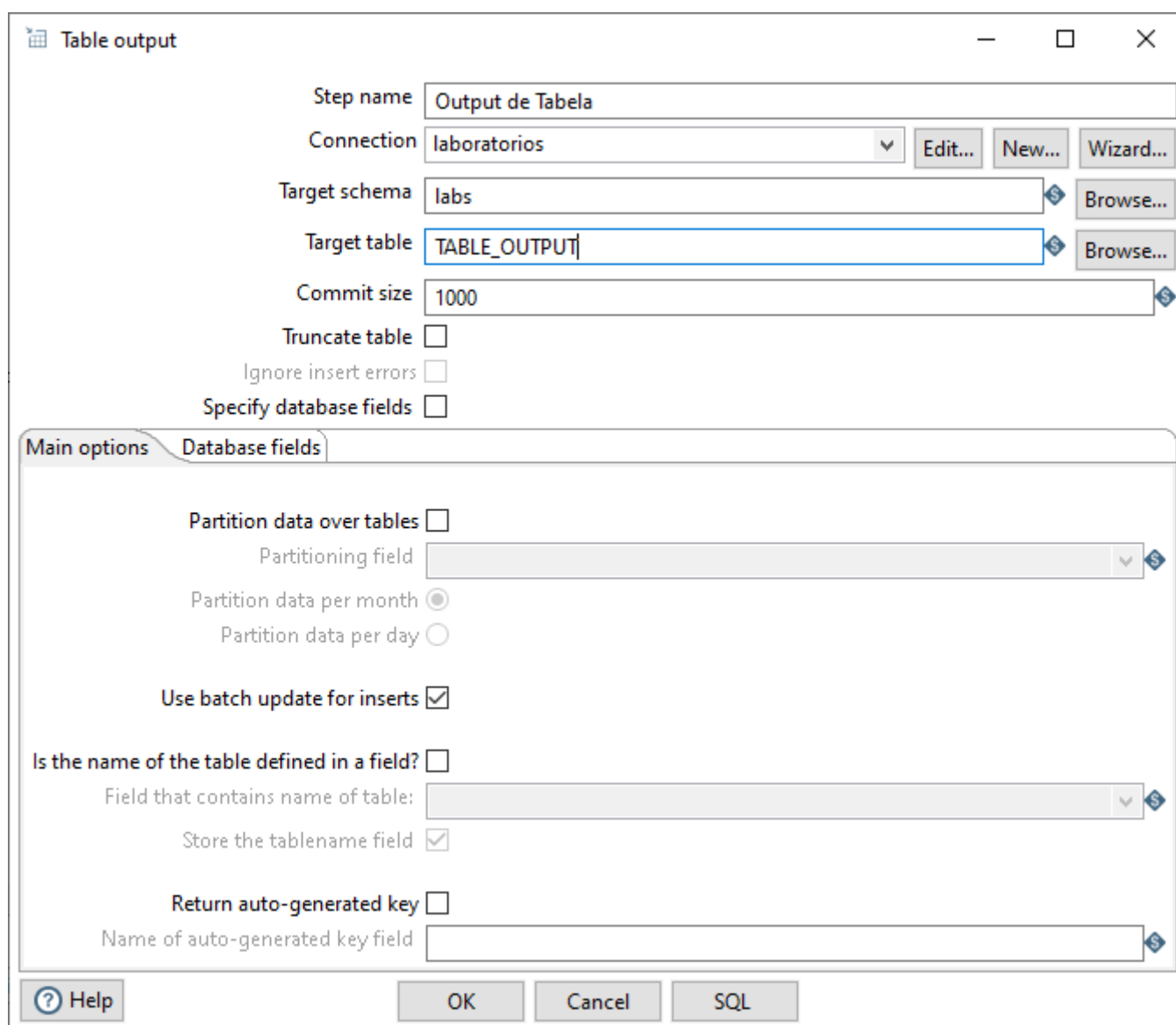


Table output

Step name: Output de Tabela

Connection: laboratorios [Edit... New... Wizard...]

Target schema: labs [Browse...]

Target table: TABLE_OUTPUT [Browse...]

Commit size: 1000

Truncate table: ☐

Ignore insert errors: ☐

Specify database fields: ☐

Main options | Database fields

Partition data over tables: ☐

Partitioning field: []

Partition data per month: ☒

Partition data per day: ☐

Use batch update for inserts: ☒

Is the name of the table defined in a field?: ☐

Field that contains name of table: []

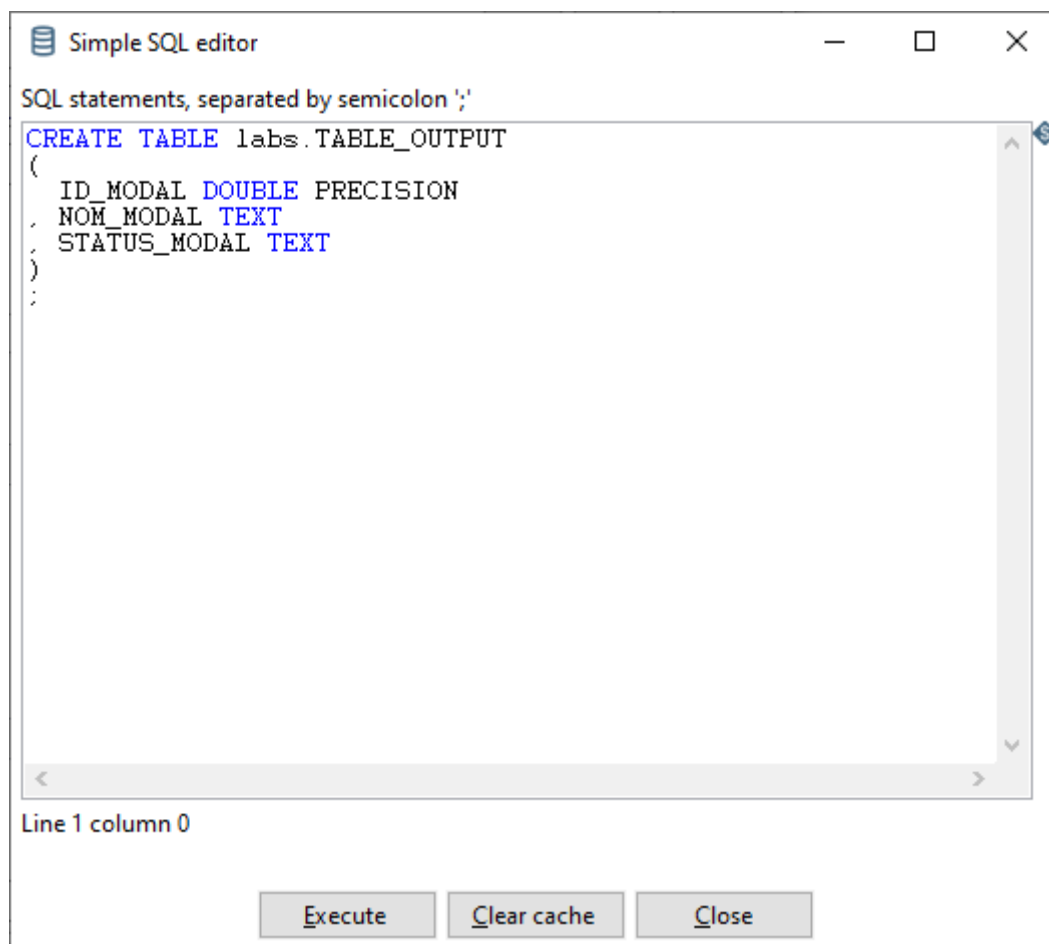
Store the tablename field: ☒

Return auto-generated key: ☐

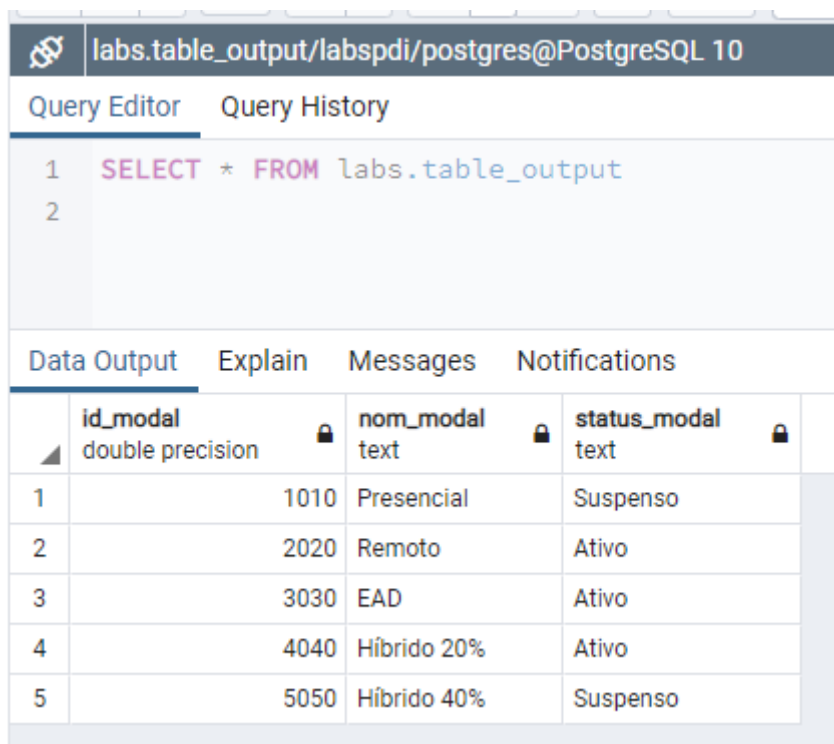
Name of auto-generated key field: []

[?] Help [OK] [Cancel] [SQL]

- 11.2** Clique em “SQL” e execute o código montando pelo PentahoDI. Esta operação irá criar a tabela no seu banco de dados, e após a execução, acesse o Postegre para localizar a tabela recém criada:



11.3 Após verificar a tabela recém criada no banco de dados, salve a transformação no PentahoDI, execute e confirme no Postgres a carga dos dados:



The screenshot shows the Pentaho Query Editor interface. The top bar indicates the connection is 'labs.table_output/labspdi/postgres@PostgreSQL 10'. The 'Query Editor' tab is active, displaying the SQL query: `SELECT * FROM labs.table_output`. Below the query editor, the 'Data Output' tab is selected, showing a table with 5 rows and 3 columns: `id_modal` (double precision), `nom_modal` (text), and `status_modal` (text). The data rows are as follows:

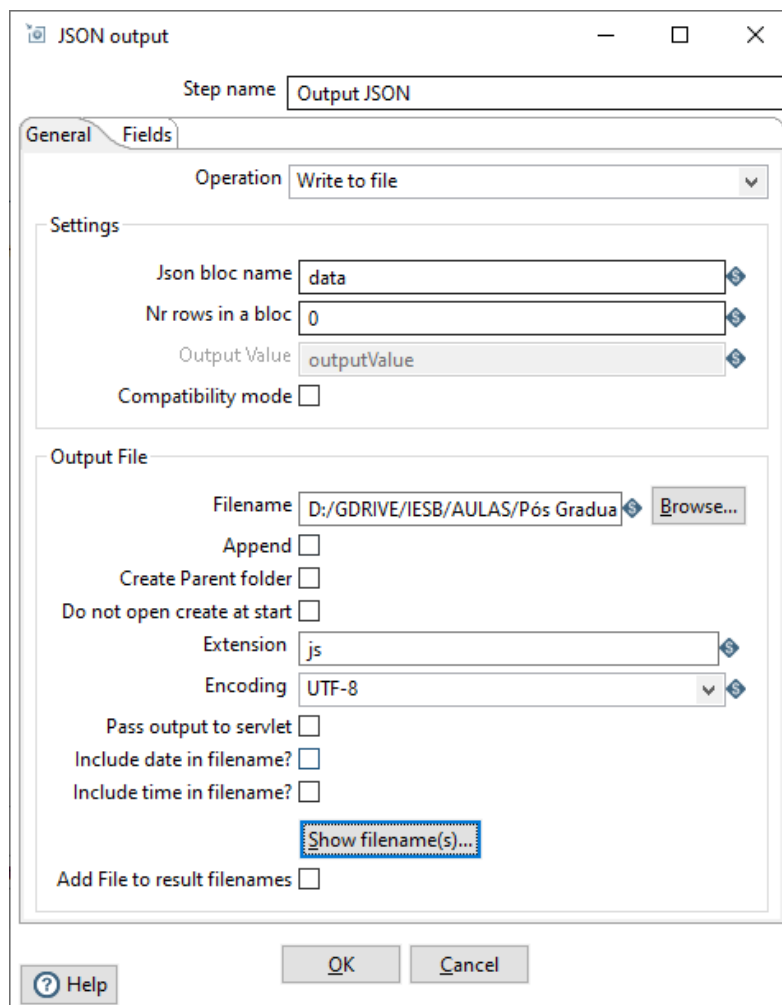
	<code>id_modal</code> double precision	<code>nom_modal</code> text	<code>status_modal</code> text
1	1010	Presencial	Suspenso
2	2020	Remoto	Ativo
3	3030	EAD	Ativo
4	4040	Híbrido 20%	Ativo
5	5050	Híbrido 40%	Suspenso

Exercício 12 – Output JSON

12.1 Crie uma nova transformação, e a salve como “12 – Output JSON”. Insira o step anterior “Dados para Output” e o step “JSON Output”, criando o HOP entre estes dois steps:



12.2 Configure o step “JSON output”, com atenção aos campos “Step name”, “Nr rows in a bloc” (deve ser marcado como zero, para que seja criado apenas um arquivo JSON de saída), “Filename” (caminho e nome do arquivo de saída), “Encoding”, localizados na aba “General”.



The screenshot shows the 'JSON output' configuration dialog box with the following settings:

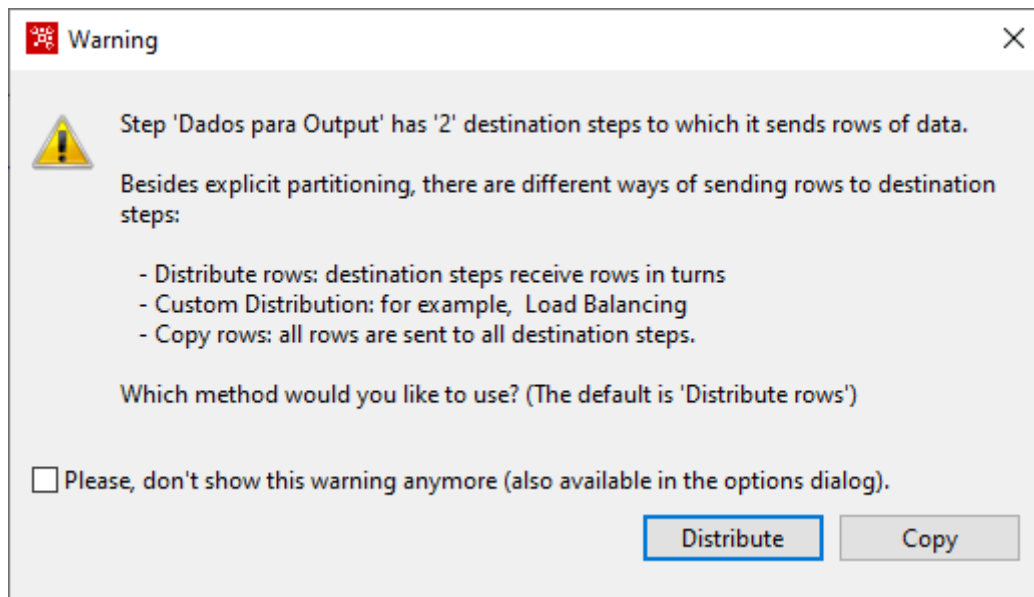
- Step name:** Output JSON
- General tab:**
 - Operation:** Write to file
 - Settings:**
 - Json bloc name:** data
 - Nr rows in a bloc:** 0
 - Output Value:** outputValue
 - Compatibility mode:** ☐
 - Output File:**
 - Filename:** D:/GDRIVE/IESB/AULAS/Pós Gradua
 - Append:** ☐
 - Create Parent folder:** ☐
 - Do not open create at start:** ☐
 - Extension:** js
 - Encoding:** UTF-8
 - Pass output to servlet:** ☐
 - Include date in filename?:** ☐
 - Include time in filename?:** ☐
 - Show filename(s):** (button)
 - Add File to result filenames:** ☐

12.3 Configure Na aba “Fields”, clique em “Get Fields”, dê “OK”, salve, execute a transformação e valide o conteúdo do arquivo gerado:

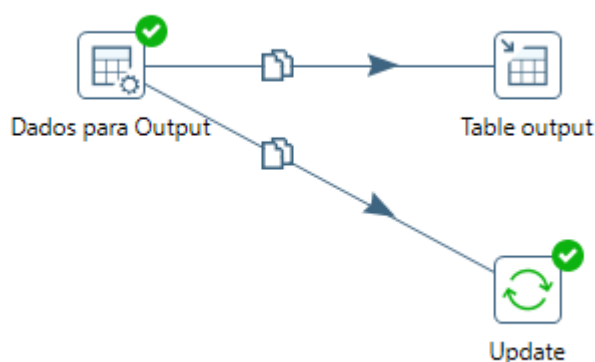
[illegible]

Exercício 13 – Output Update

13.1 Crie uma nova transformação, e a salve como “13 – Output Update”. Insira os steps “Dados para Output”, “Table Output” e “Update”. Crie inicialmente um HOP entre os steps de entrada e saída em tabela. Em seguida, crie um segundo HOP entre os dados de entrada e o step “Update”. Na janela que se surgirá, selecione “Copy”:

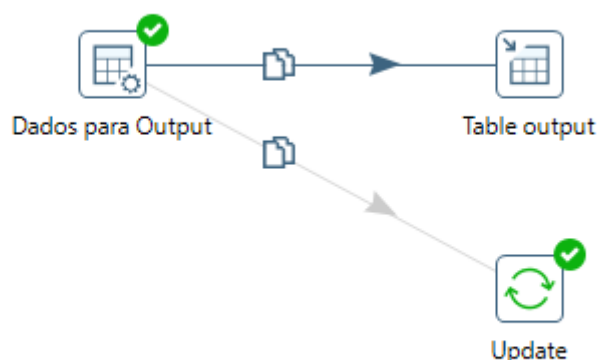


13.2 A transformação deverá ficar como a seguir:

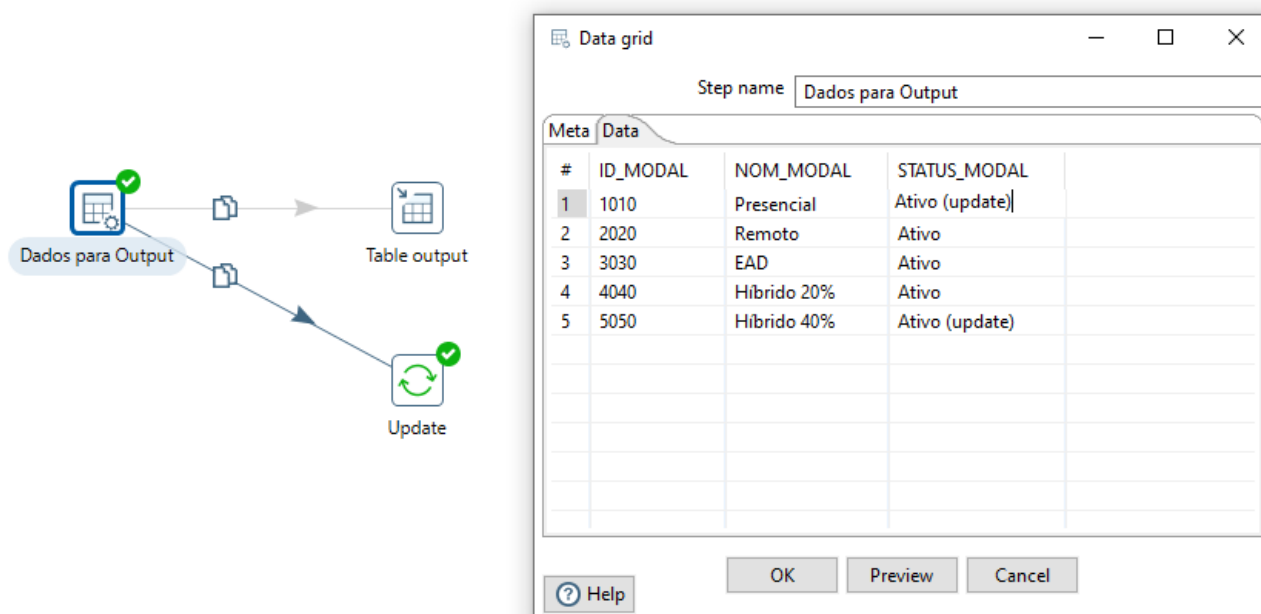


Obs.: Ao se selecionar “Distribute”, os dados serão literalmente distribuídos entre cada step, cada step com uma parte única dos dados. Já na opção “Copy”, a totalidade dos dados é enviada a cada step.

13.3 A Em seguida, configure o step “Table output para criar uma nova tabela, e a carregue com os dados tais quais existem no step data grid “Dados para Output”. Antes de executar a transformação, dê um único clique no HOP entre os dados de entrada e o step de update, para que este seja desabilitado para a primeira execução:



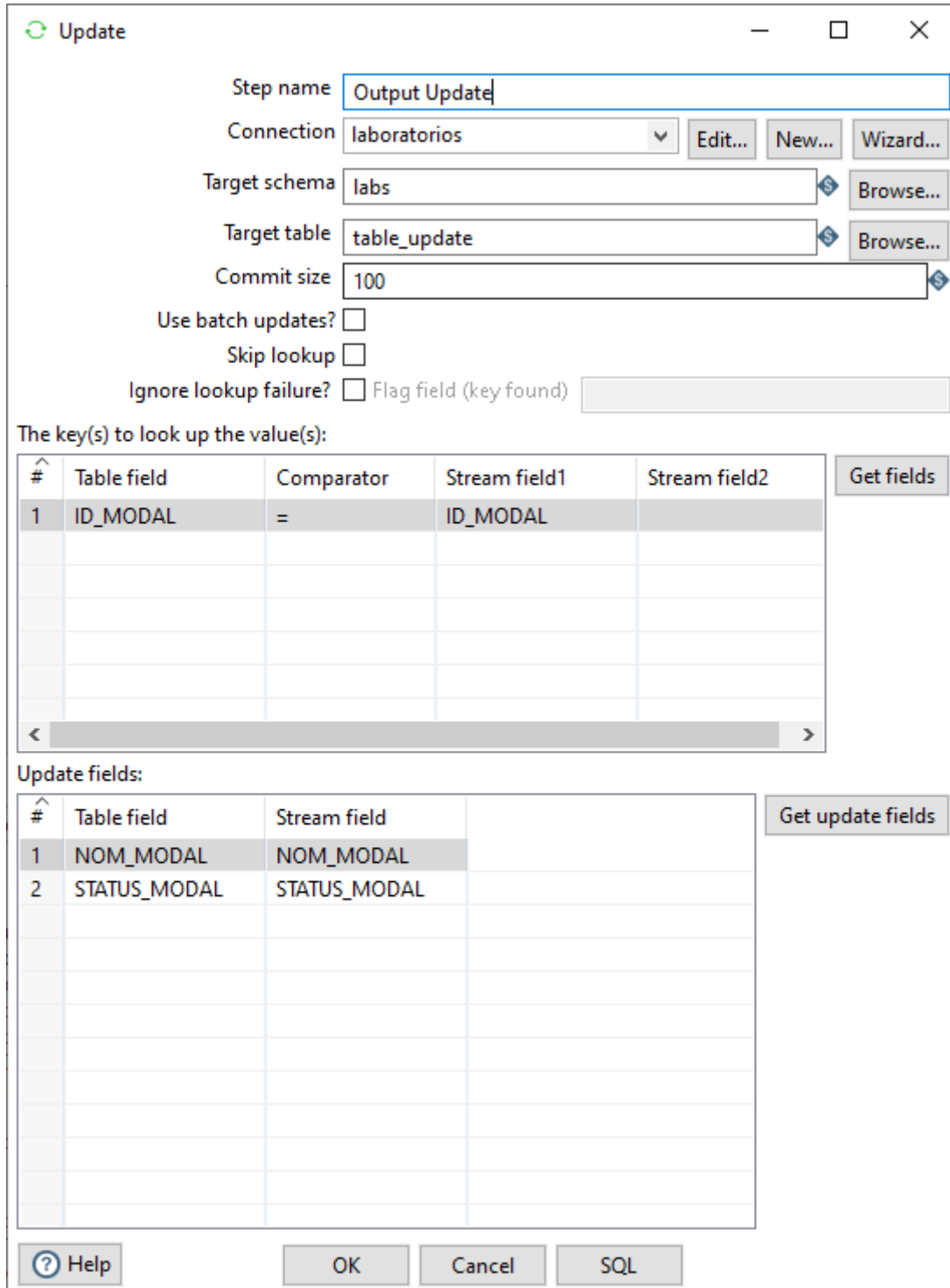
13.4 Verifique a carga de dados na tabela de destino. Altere o conteúdo do step “Dados para Output”, desabilite o HOP entre o step de entrada e a saída de tabela, reabilitando o HOP entre o step de entrada e o step “Update”:



O diagrama de fluxo de dados ETL mostra o step 'Dados para Output' conectado ao step 'Table output' e ao step 'Update'. A janela 'Data grid' está aberta, mostrando os dados carregados no step 'Dados para Output'.

Meta	Data		
#	ID_MODAL	NOM_MODAL	STATUS_MODAL
1	1010	Presencial	Ativo (update)
2	2020	Remoto	Ativo
3	3030	EAD	Ativo
4	4040	Híbrido 20%	Ativo
5	5050	Híbrido 40%	Ativo (update)

13.5 Antes de executar, é necessário configurar o step “Update”. Altere o nome do step em “Step name”, selecione a conexão em “Connection”, a tabela em “Target table”. Feito isto, é preciso indicar a chave de lookup para os registros. Clique em “Get fields”, após a leitura dos campos, mantenha apenas a chave da tabela, no caso, “ID_MODAL”, apagando os demais campos. Em “Update fields”, serão indicados os campos que poderão sofrer alterações, portanto, o campo “ID_MODAL” deve ser removido, mantendo apenas os campos “NOM_MODAL” e “STATUS_MODAL”. Salve e execute a transformação:



Update

Step name:

Connection:

Target schema:

Target table:

Commit size:

Use batch updates? ☐

Skip lookup ☐

Ignore lookup failure? ☐ Flag field (key found)

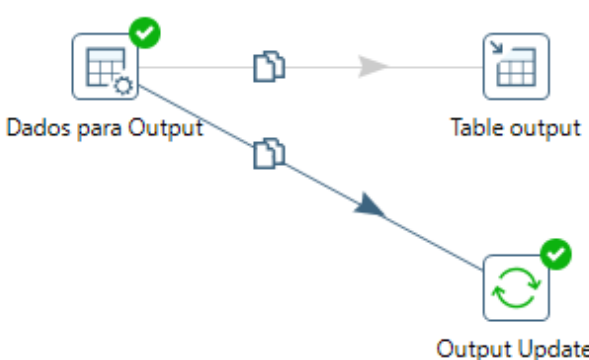
The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1	ID_MODAL	=	ID_MODAL	

Update fields:

#	Table field	Stream field
1	NOM_MODAL	NOM_MODAL
2	STATUS_MODAL	STATUS_MODAL

13.6 Em “Execution Results”, vá em “Preview data” e localize as alterações realizadas:



The diagram shows an ETL workflow. A source node labeled 'Dados para Output' (with a green checkmark) connects to two target nodes: 'Table output' and 'Output Update' (both with green checkmarks). The 'Output Update' node has a circular arrow icon, indicating an update operation.

Execution Results

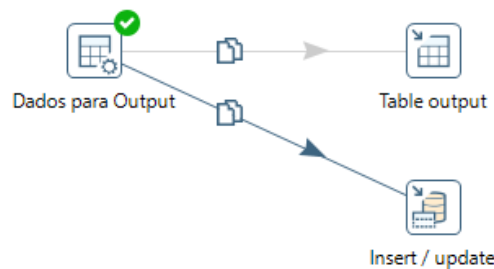
Logging | Execution History | Step Metrics | Performance Graph | Metrics | Preview data

☒ First rows ☐ Last rows ☐ Off

#	ID_MODAL	NOM_MODAL	STATUS_MODAL
1	1010	Presencial	Ativo (update)
2	2020	Remoto	Ativo
3	3030	EAD	Ativo
4	4040	Híbrido 20%	Ativo
5	5050	Híbrido 40%	Ativo (update)

Exercício 14 – Output Insert/Update

14.1 Abra novamente a transformação “13 – Output Update” e a salve como “14 – Output Insert_Update”. Apague o step “Update” e insira o step “Insert / Update” em seu lugar:



14.2 Volte manualmente o conteúdo do step “Dados para Output” para seus dados originais, desabilite o HOP para o step “Insert / Update” e carregue a tabela de destino. Em seguida, configure o step “Insert / Update”:

Insert / update

Step name: Output Insert/Update

Connection: laboratorios

Target schema: labs

Target table: table_update

Commit size: 100

Don't perform any updates: ☐

The key(s) to look up the value(s):

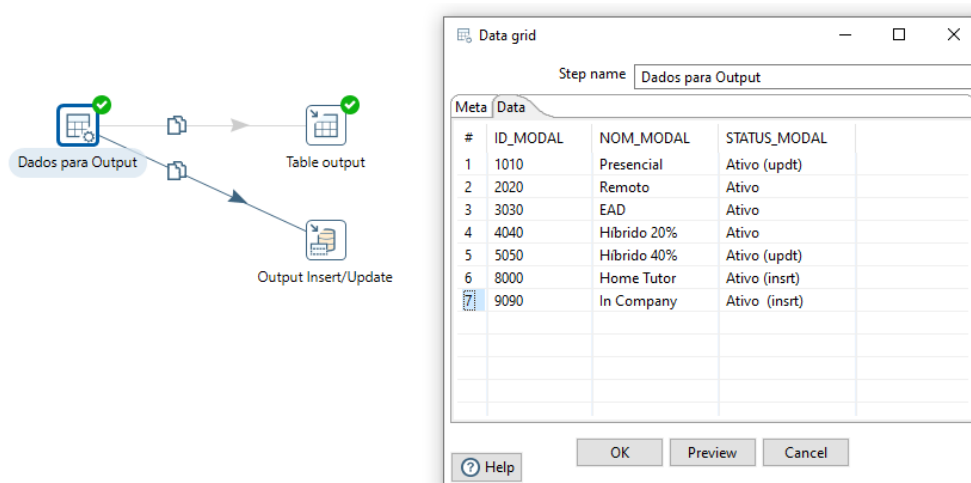
#	Table field	Comparator	Stream field1	Stream field
1	ID_MODAL	=	ID_MODAL	

Update fields:

#	Table field	Stream field	Update
1	NOM_MODAL	NOM_MODAL	Y
2	STATUS_MODAL	STATUS_MODAL	Y

Buttons: Help, OK, Cancel, SQL, Get fields, Get update fields, Edit mapping

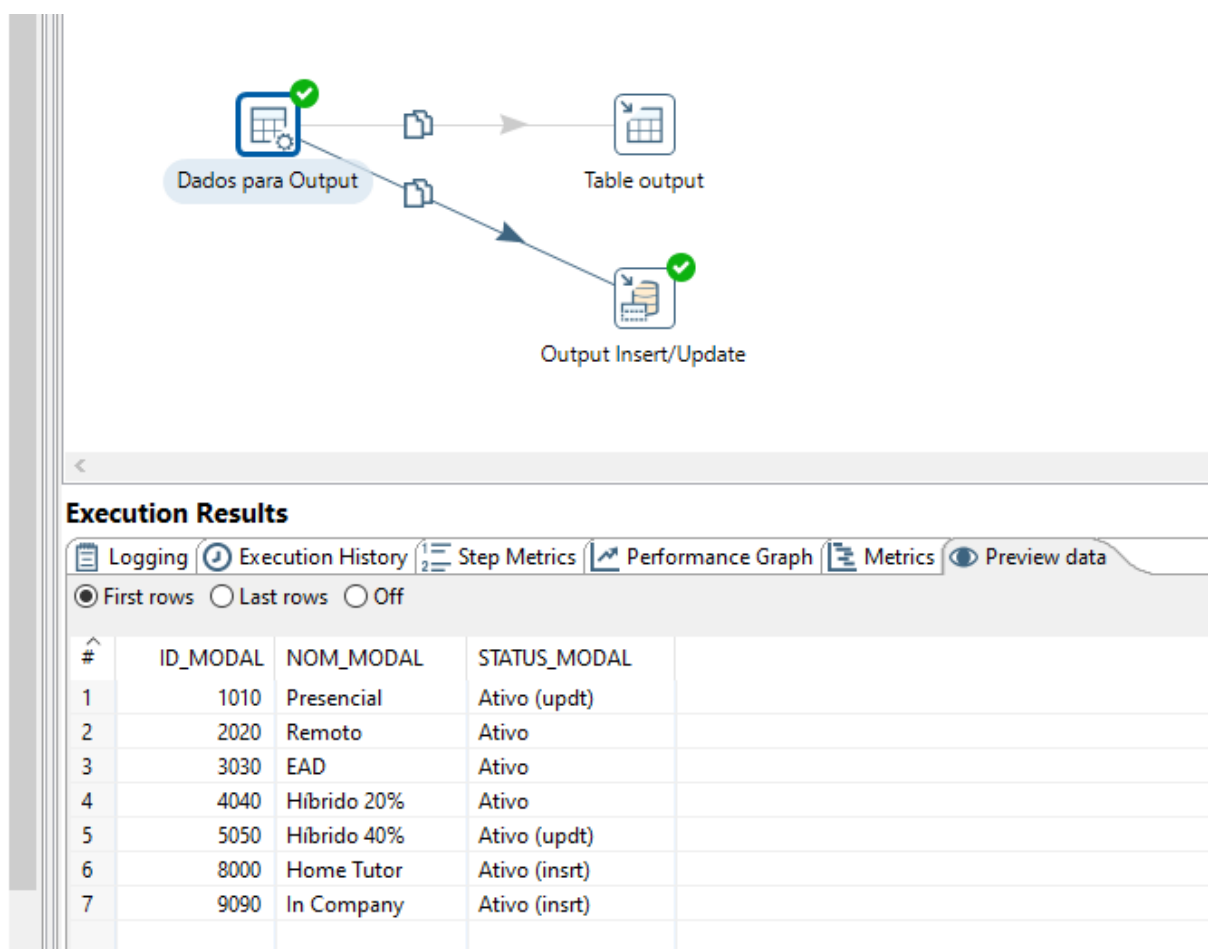
14.3 Edite o conteúdo do Data Grid para que represente duas alterações em suas linhas e adicione também mais duas linhas:



O diagrama de fluxo ETL mostra a transformação 'Dados para Output' sendo enviada para 'Table output' e 'Output Insert/Update'. A janela 'Data grid' para o passo 'Dados para Output' contém a seguinte tabela:

#	ID_MODAL	NOM_MODAL	STATUS_MODAL
1	1010	Presencial	Ativo (updt)
2	2020	Remoto	Ativo
3	3030	EAD	Ativo
4	4040	Híbrido 20%	Ativo
5	5050	Híbrido 40%	Ativo (updt)
6	8000	Home Tutor	Ativo (insrt)
7	9090	In Company	Ativo (insrt)

14.4 Salve a transformação, execute o processo e na guia “Execution Results”, na aba “Preview data”, verifique se houve duas alterações no conteúdo dos dados e duas novas linhas:



O diagrama de fluxo ETL mostra a transformação 'Dados para Output' sendo enviada para 'Table output' e 'Output Insert/Update'. A aba 'Preview data' da guia 'Execution Results' mostra a seguinte tabela:

#	ID_MODAL	NOM_MODAL	STATUS_MODAL
1	1010	Presencial	Ativo (updt)
2	2020	Remoto	Ativo
3	3030	EAD	Ativo
4	4040	Híbrido 20%	Ativo
5	5050	Híbrido 40%	Ativo (updt)
6	8000	Home Tutor	Ativo (insrt)
7	9090	In Company	Ativo (insrt)