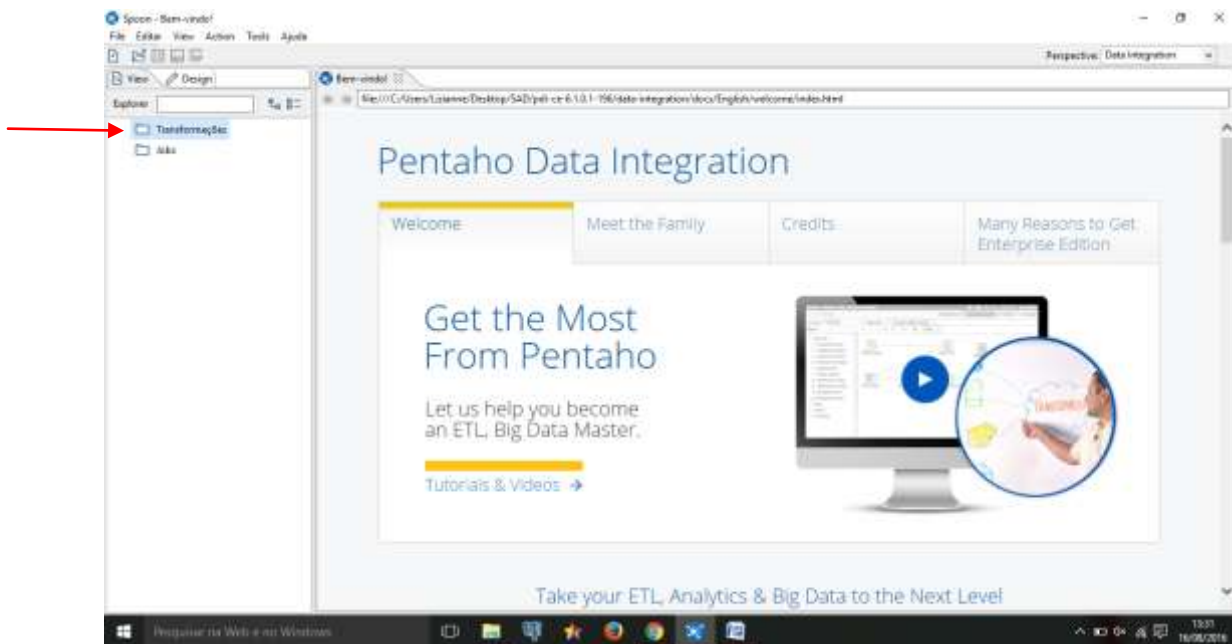


TUTORIAL: Como criar um pequeno ETL com o PDI Pentaho, utilizando informações advindas dos sistemas do Governo para servir de carga.

Passo 1: Baixe o arquivo de dados no site do Governo [<http://dados.gov.br/> ].

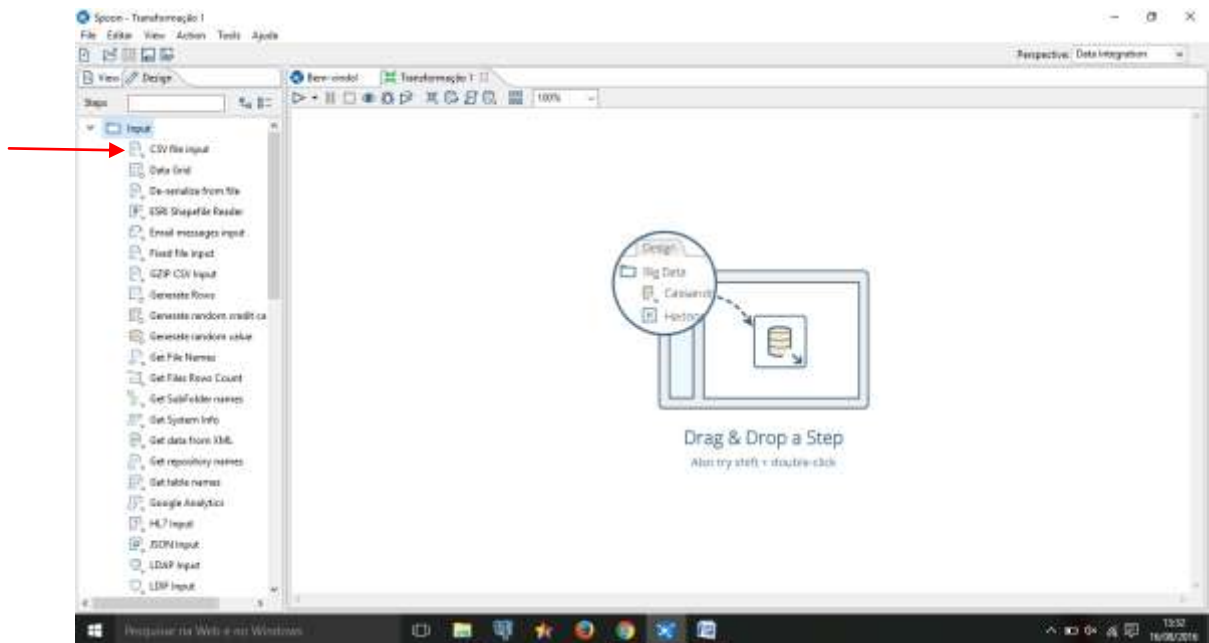
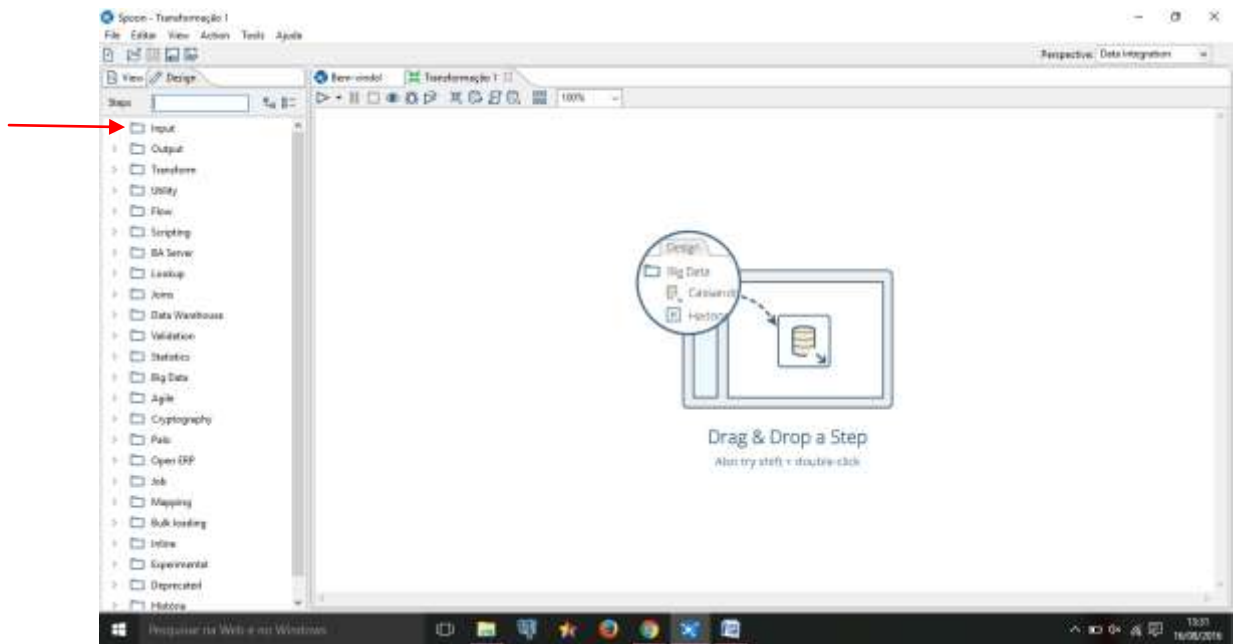
Passo 2: Execute o arquivo Spoon.bat do PDI Pentaho . A seguinte tela aparecerá:

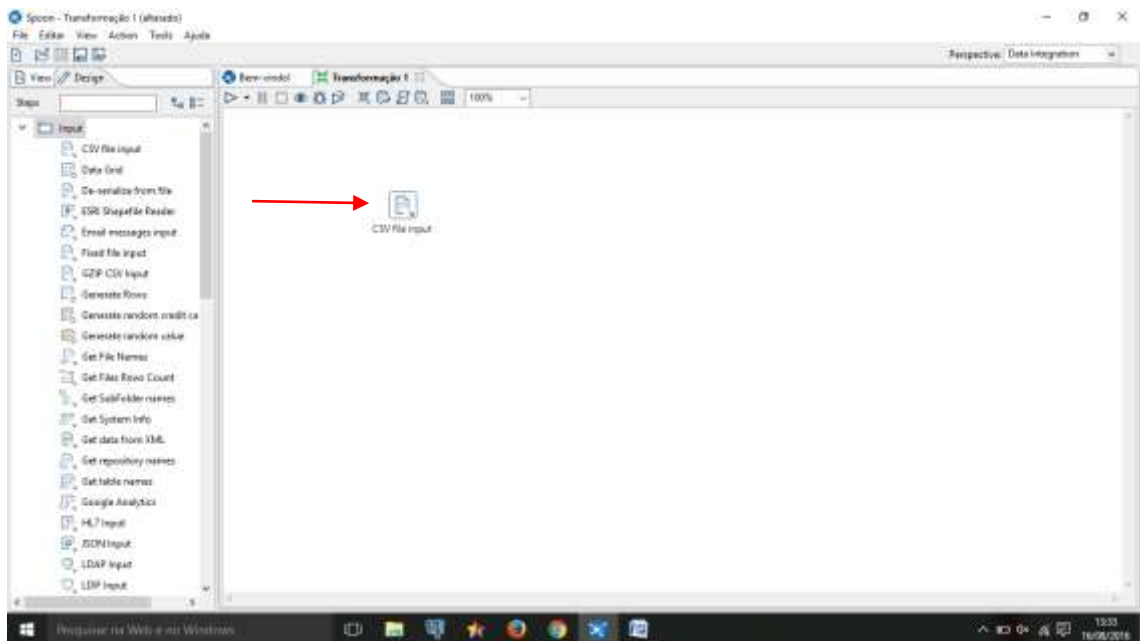


Clique em Transformações para criar uma nova Transformação.

Passo 3: Agora escolha a Entrada ("Input") de acordo com tipo do seu arquivo.

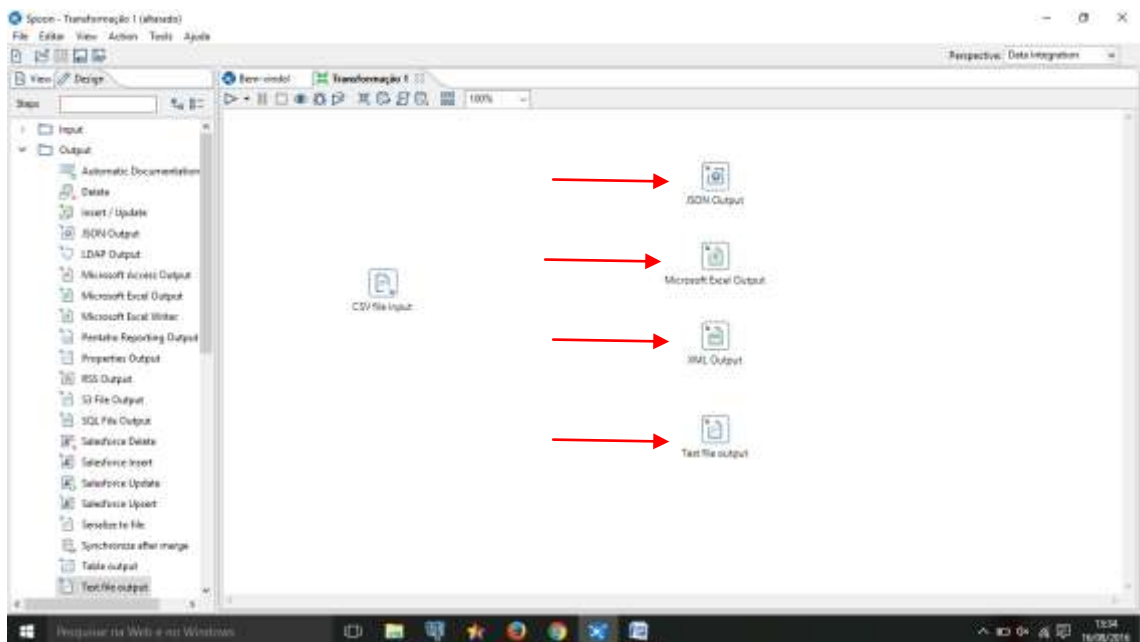
Neste caso, poderá ser .xls ou .csv, faremos com o .csv, clique duas vezes ou arraste o arquivo para o local indicado.



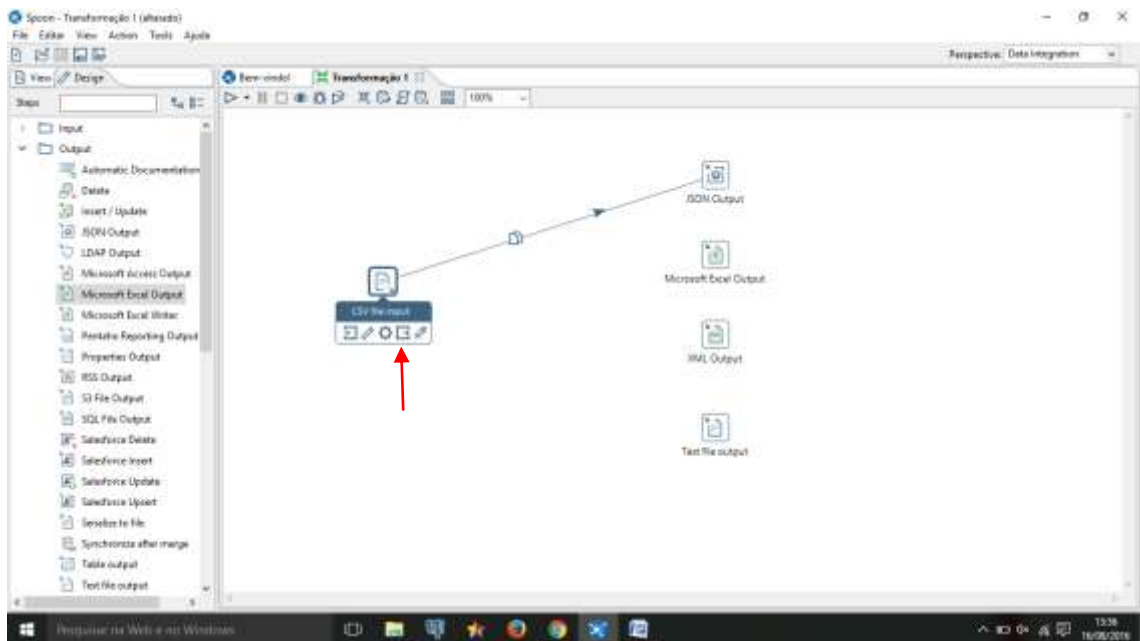


Passo 4 : Agora, escolha os arquivos de saída ("Output").

Neste caso, as saídas serão do tipo JSON, XLS, TXT e XML.

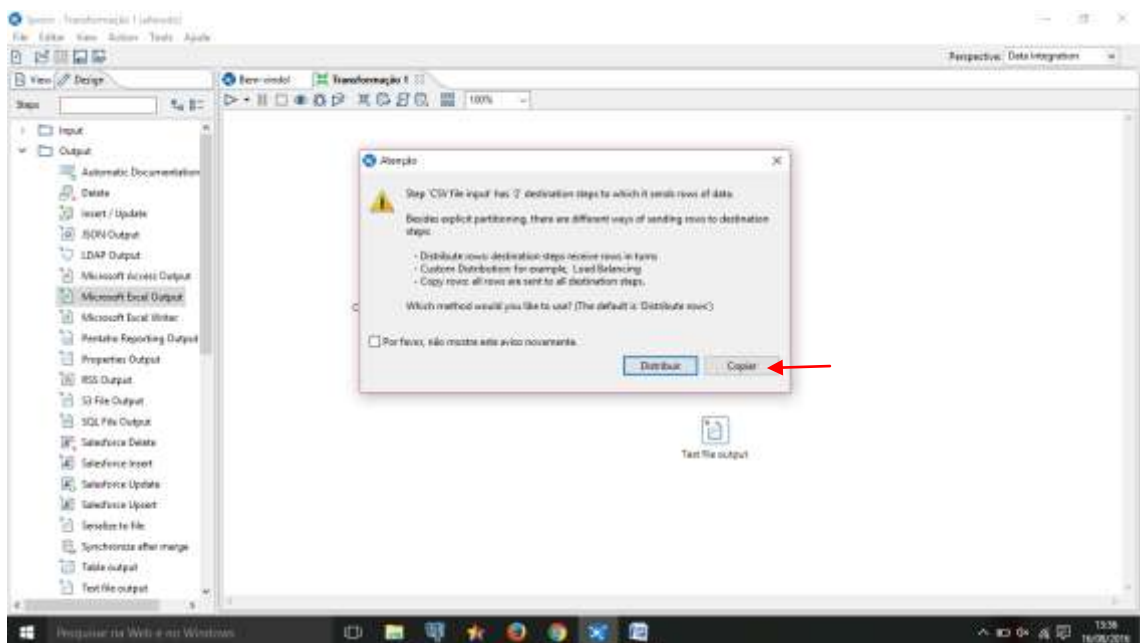


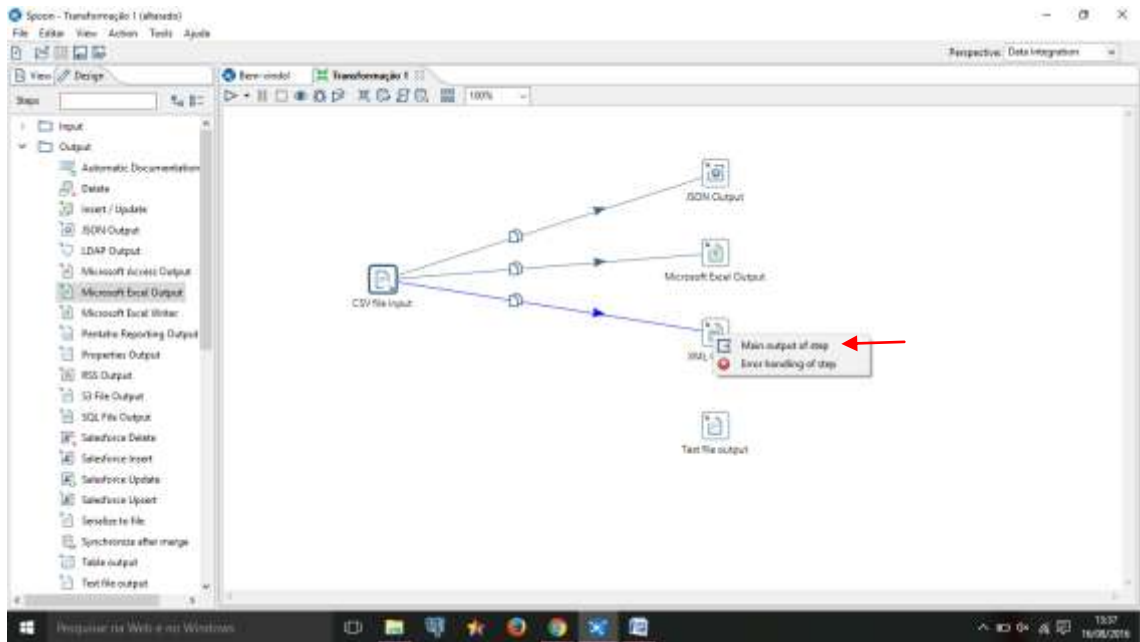
Passo 5: Agora, clique no arquivo de ENTRADA e faça as ligações para as respectivas SAÍDAS.



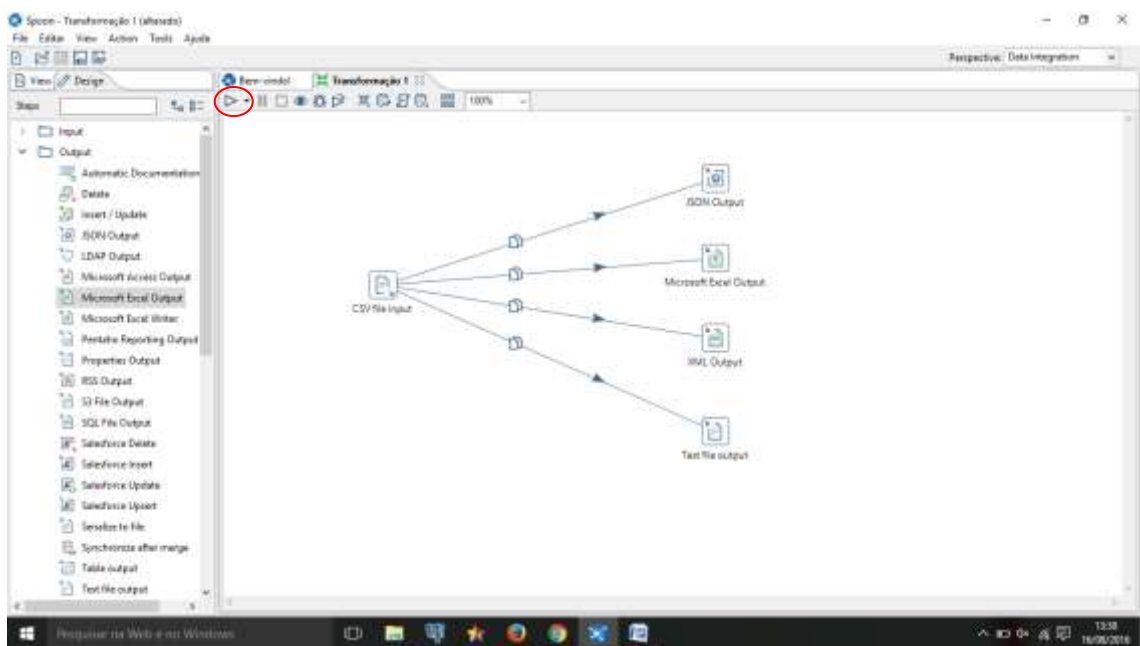
Passo 6: Uma mensagem aparecerá a respeito de: Se os arquivos devem ser distribuídos ou copiados.

Neste caso, optamos por copiar todas as linhas da entrada para todas as saídas, clicando em COPIAR.

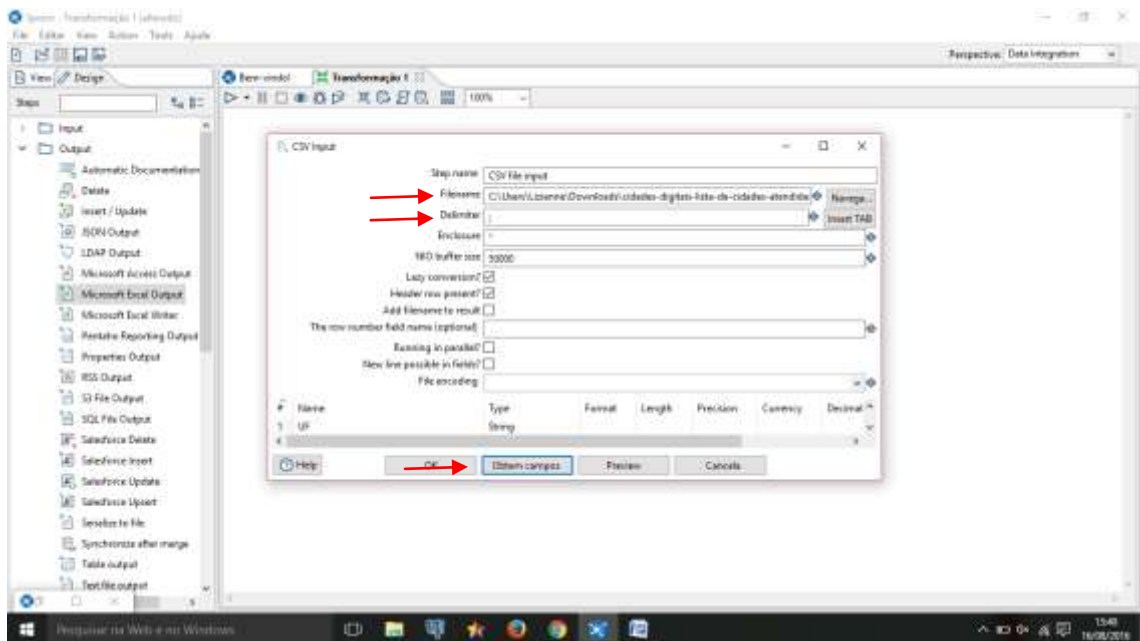




Passo 7 : Agora clique em PLAY, para executar e em seguida "RUN":

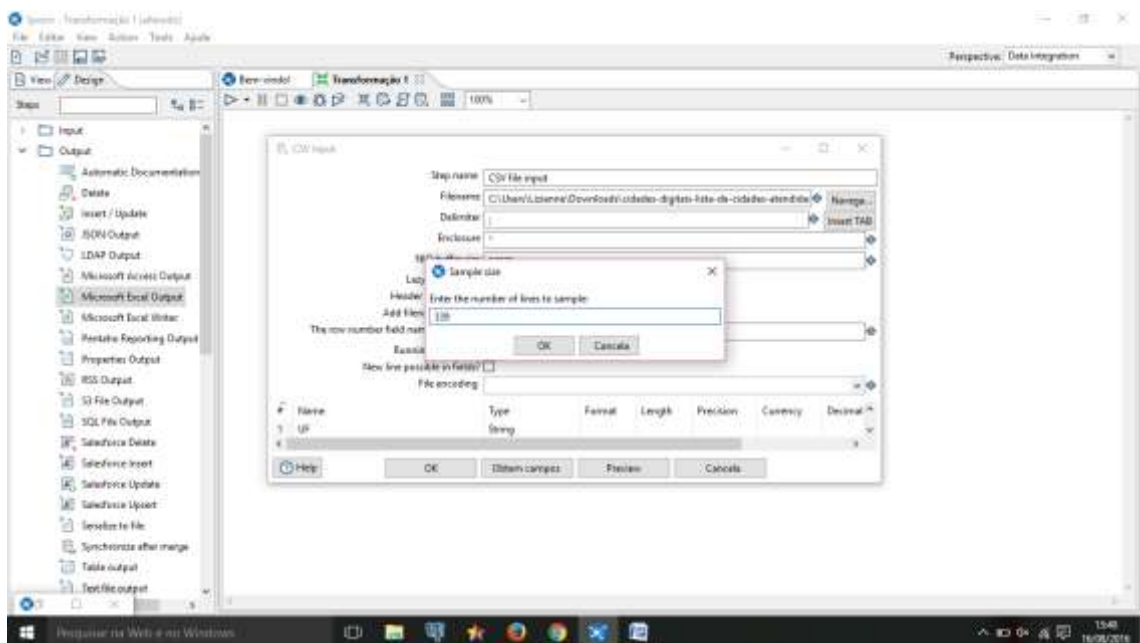


A seguinte janela abrirá:

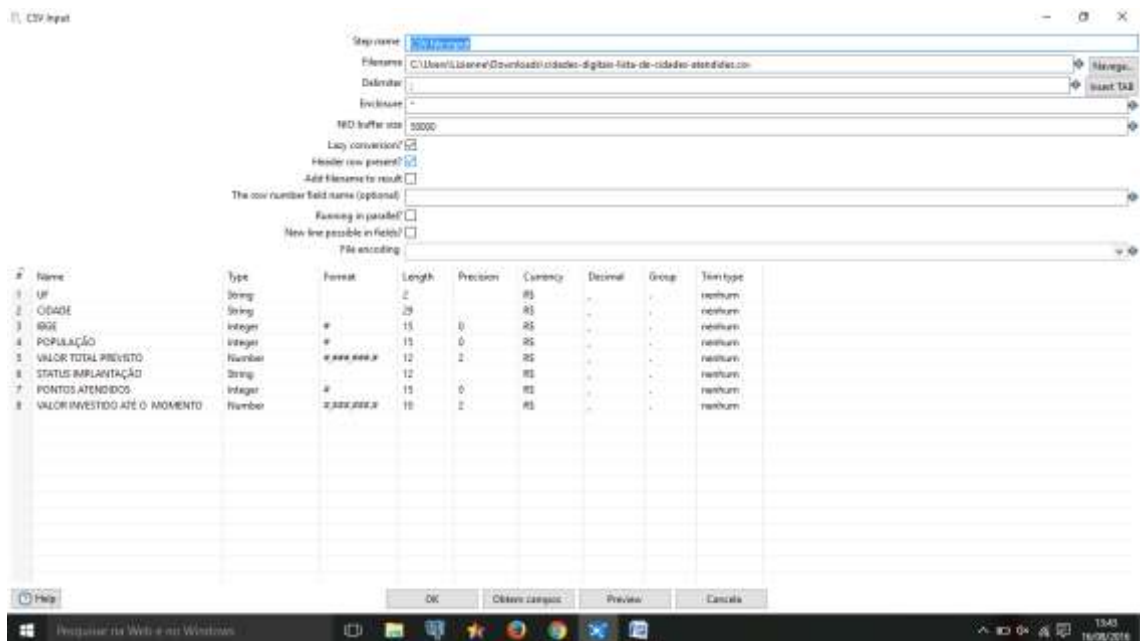
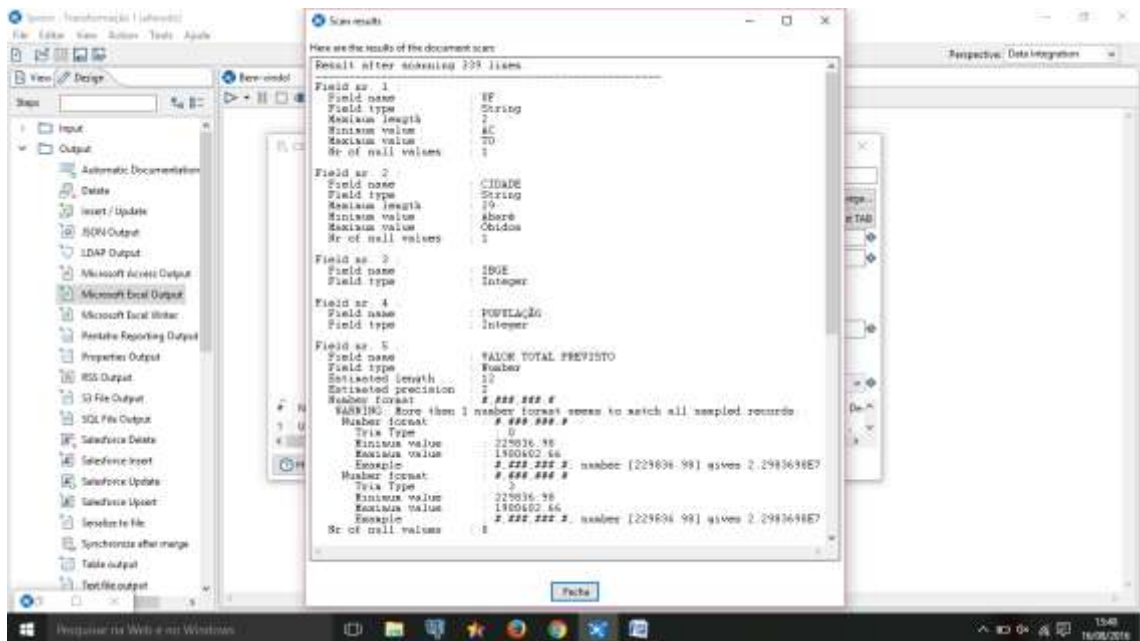


Coloque selecione o arquivo com os dados (Filename) e digite o delimitador/separador entre os dados do arquivo. Agora clique em "OBTEM CAMPOS" para visualizar.

Passo 8: Digite o número de linhas do arquivo de dados:



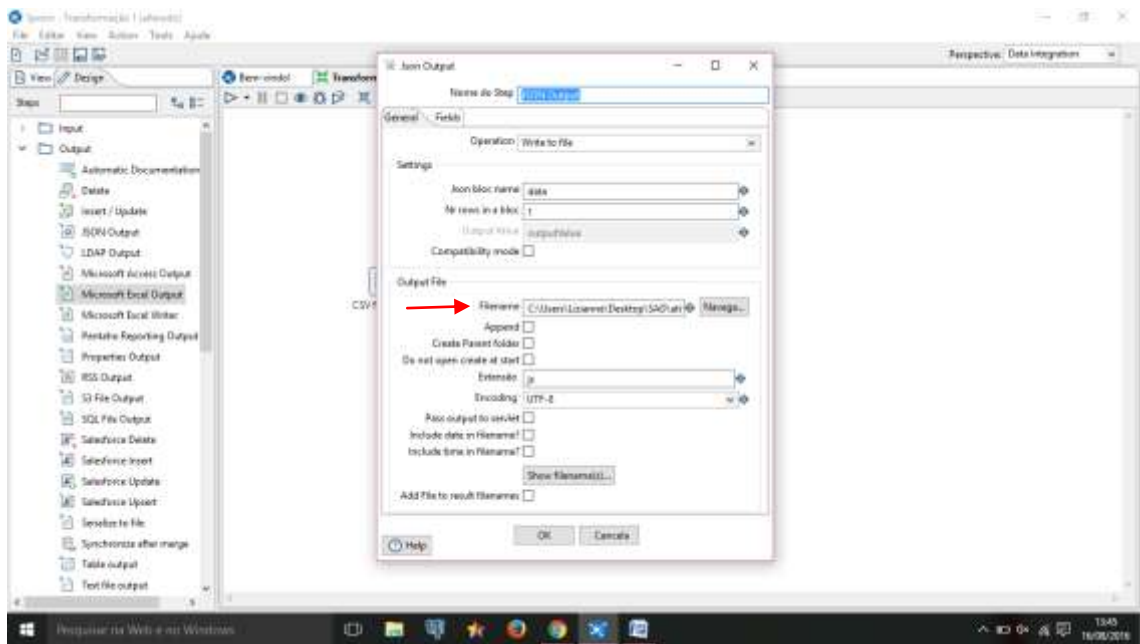
Passo 9 : O resultado.



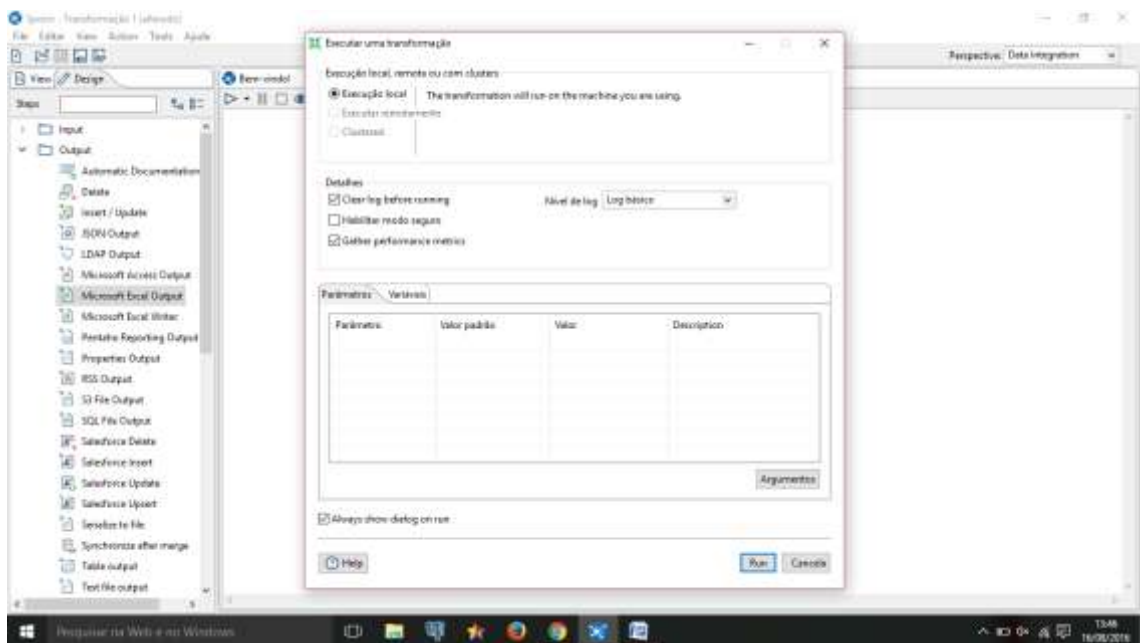
Clique OK para salvar e sair.

Passo 10: Agora, clique duas vezes na saída ( o mesmo processo será feito para TODAS as saídas).

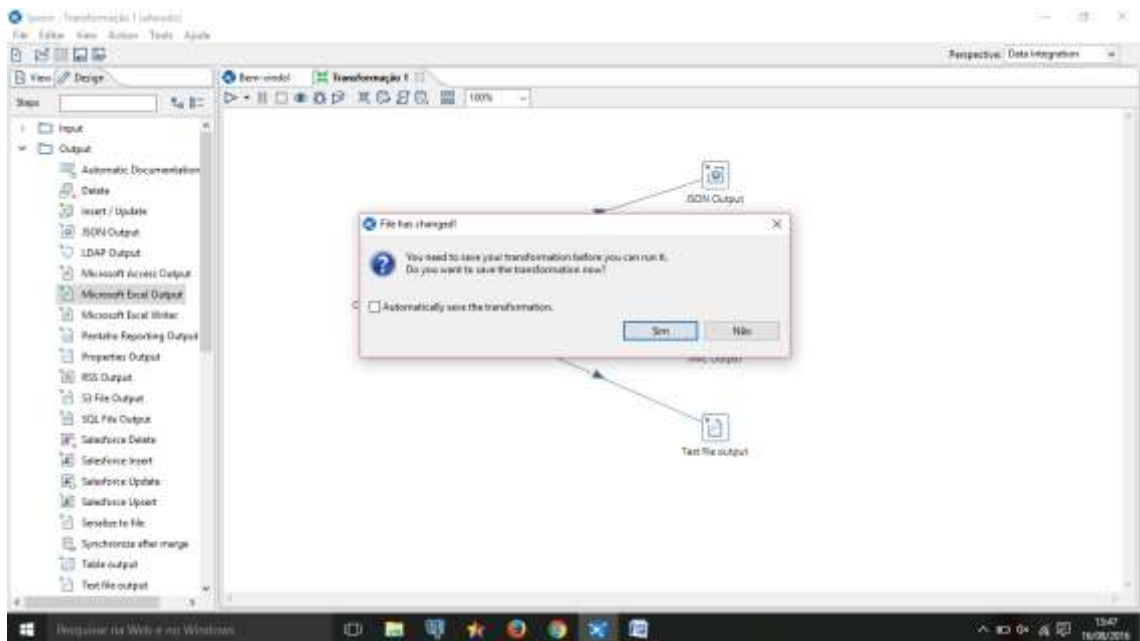
Em "FILENAME" indique o local para salvar o arquivo e dê um nome para o mesmo.



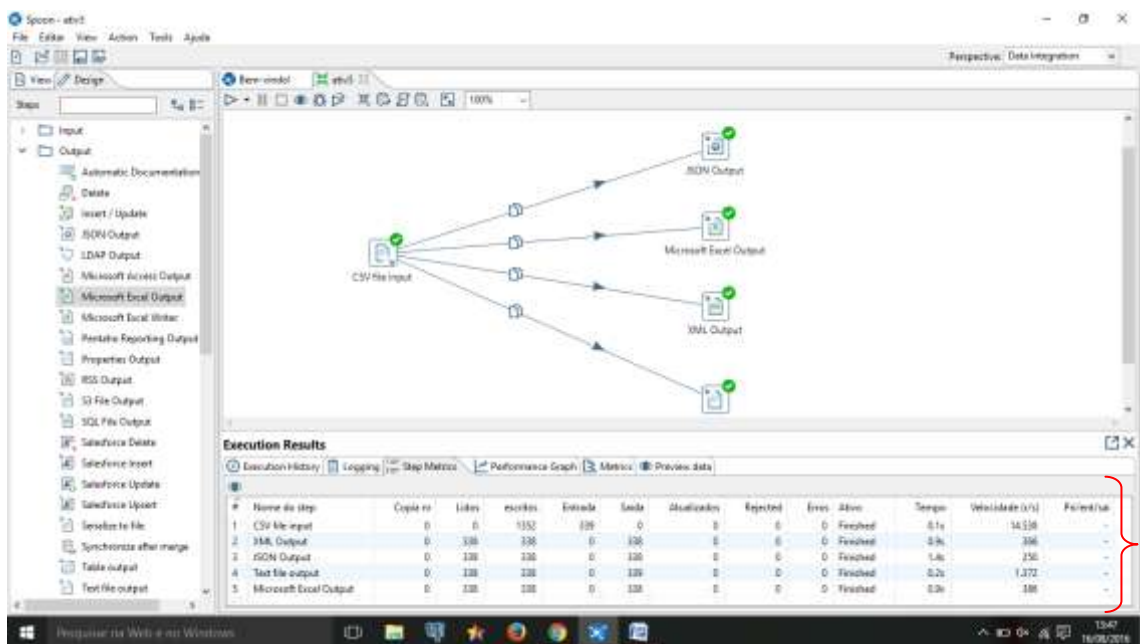
Confirme e prossiga.







Quando os arquivos forem executados sem nenhum erro, "as entradas e saídas" aparecerão com o ícone verde como na figura abaixo.



Resultado da execução.

Para confirmar os resultados, compare os arquivos, verifique a quantidade de entrada de saída, por exemplo.