

## Atividade 5

Atividade voltada para as noções básicas de ETL com o PDI Pentaho.

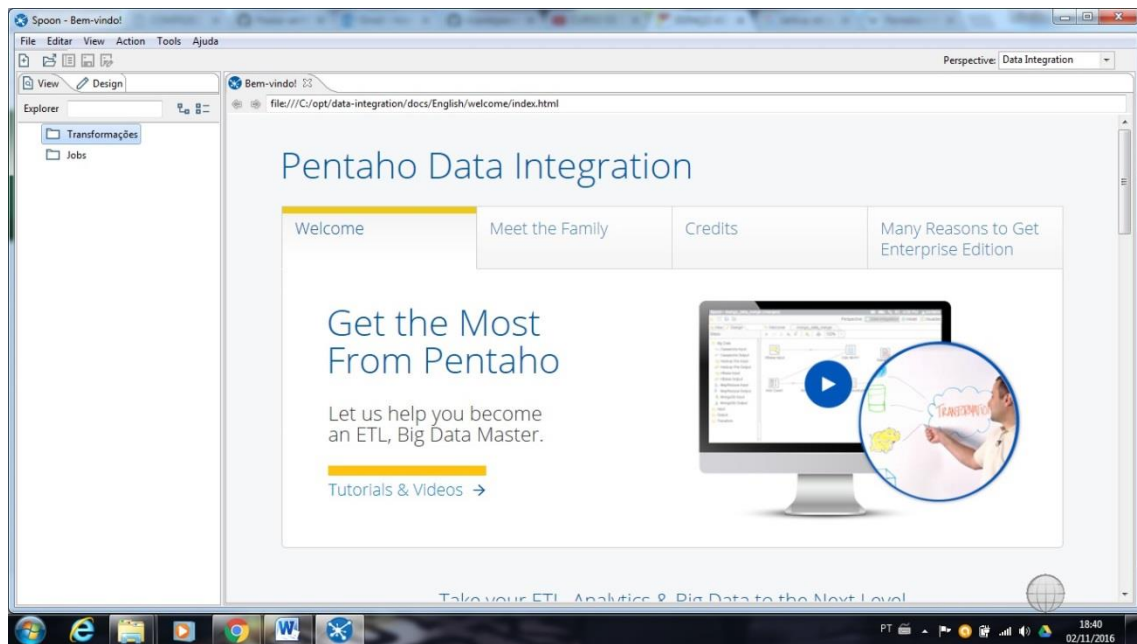
Crie uma pequena documentação/tutorial de como realizar os seguintes tópicos (tire snapshots das telas e escreva de forma explicativa):

- A partir do site [<http://dados.gov.br/> ], utilize as informações advindas dos sistemas do Governo para servir de carga
- Inicie o PDI Pentaho e crie um pequeno ETL:

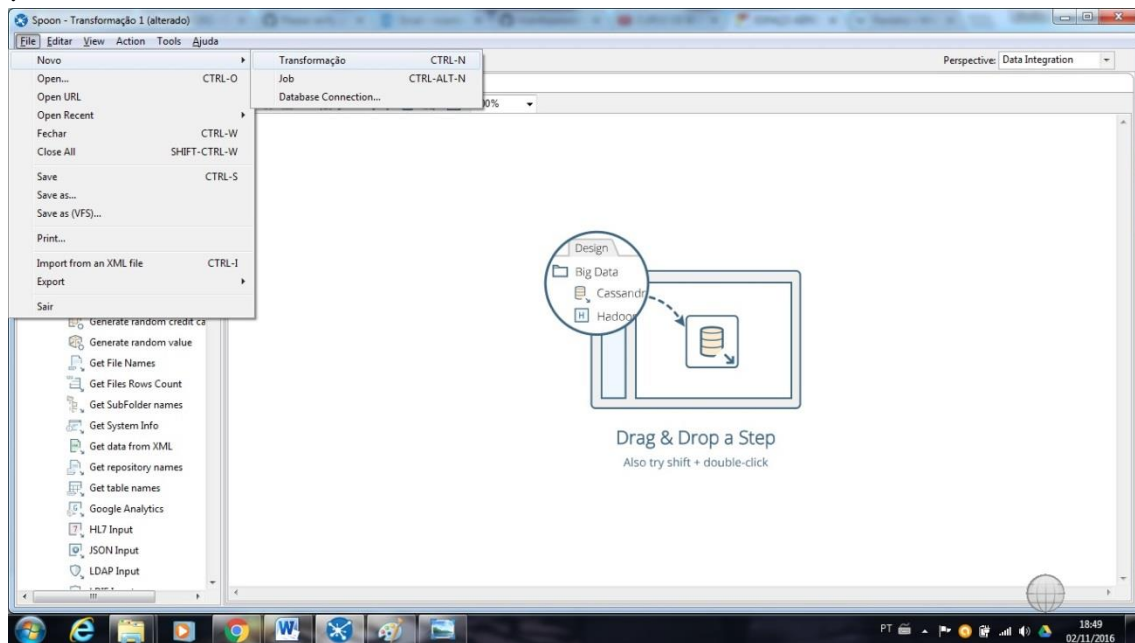
ENTRADA: XLS ou CSV

SAÍDAS: JSON, XLS, TXT e XML

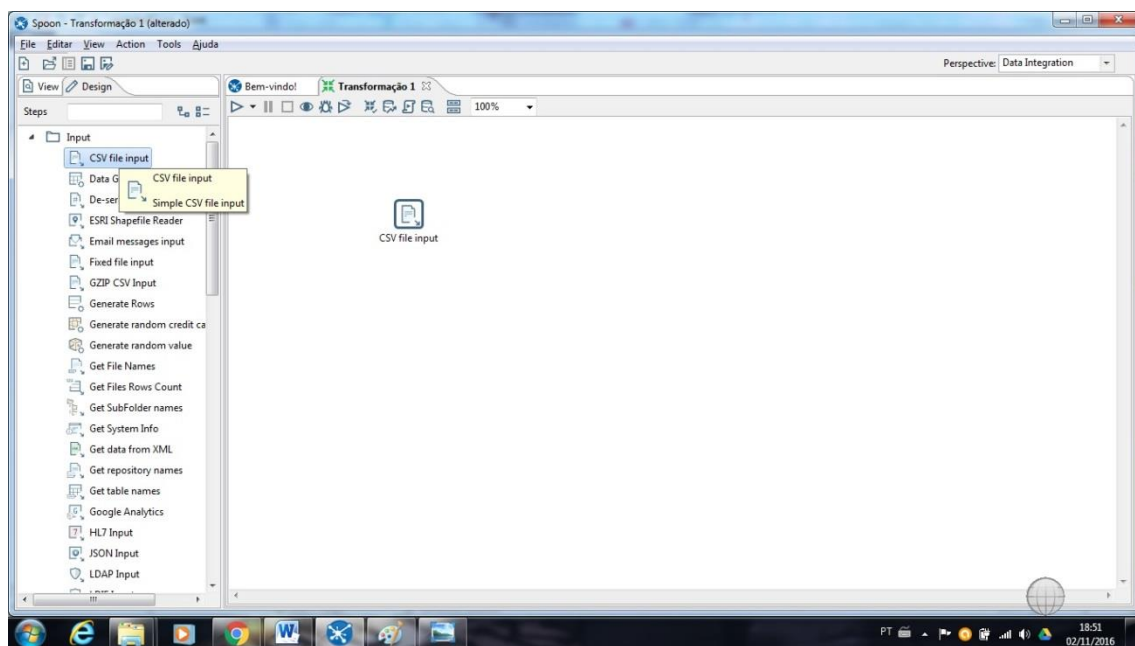
Partimos da execução do PDI(Pentaho Data Integration), também conhecido como Kettle:



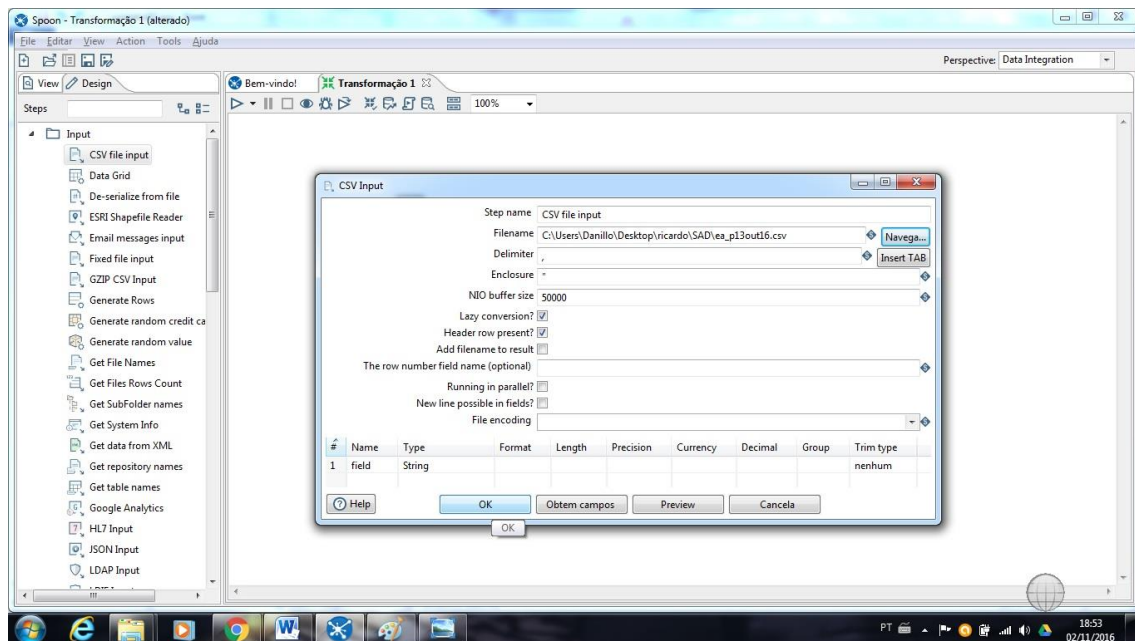
Para fazer uma transformação de dados, seguimos da seguinte forma, clicando no botão “File”, em seguida em “Novo” e por fim, clicamos em “Transformação”. Segue Imagem:



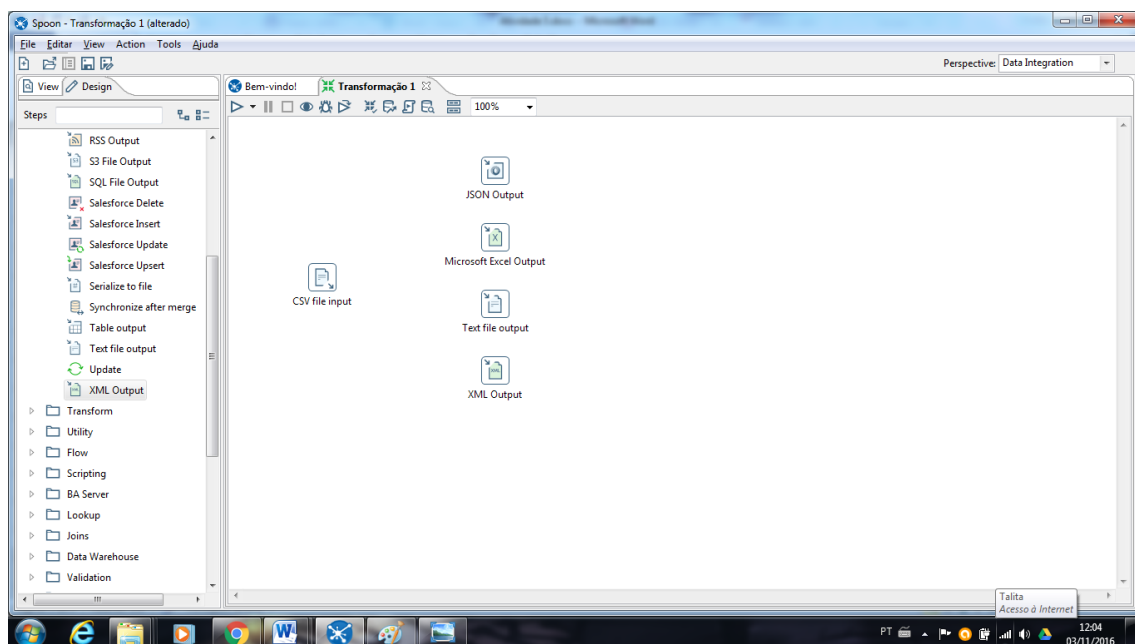
Na tela demonstrada a seguir clicamos em “Input”, aparece uma lista com diversas opções. Clicando em “CSV file Input” mantemos o botão do mouse pressionado e arrastamos até a tela ao lado. Assim criamos a interface de entrada.




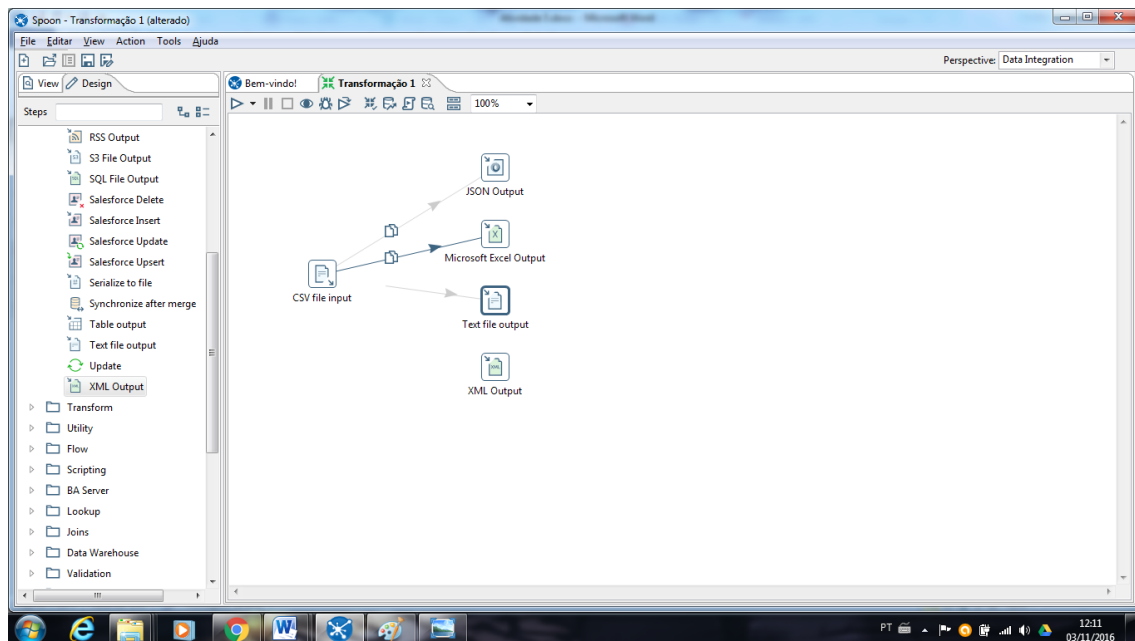
Em seguida, damos um duplo clique no ícone criado, abrindo a janela mostrada a seguir. Nesta janela, no campo “Filename”, clicamos no botão “Navega” para buscar o arquivo desejado. Para concluir, clicamos em “OK”

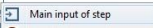


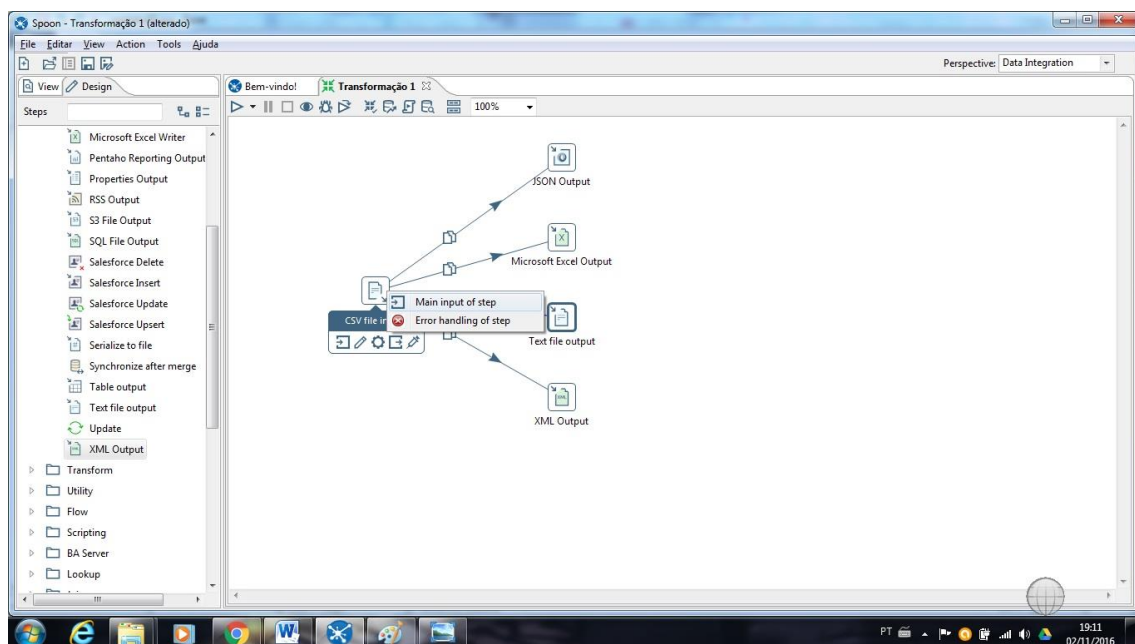
Em seguida, de forma bem parecida com os passos anteriores, vamos escolher as interfaces de Saída. Do lado esquerdo da tela, logo abaixo de “Input”, clicamos na opção “Output”, após a exibição das opções de saída, clicamos e arrastamos as escolhidas para a tela do lado direito.



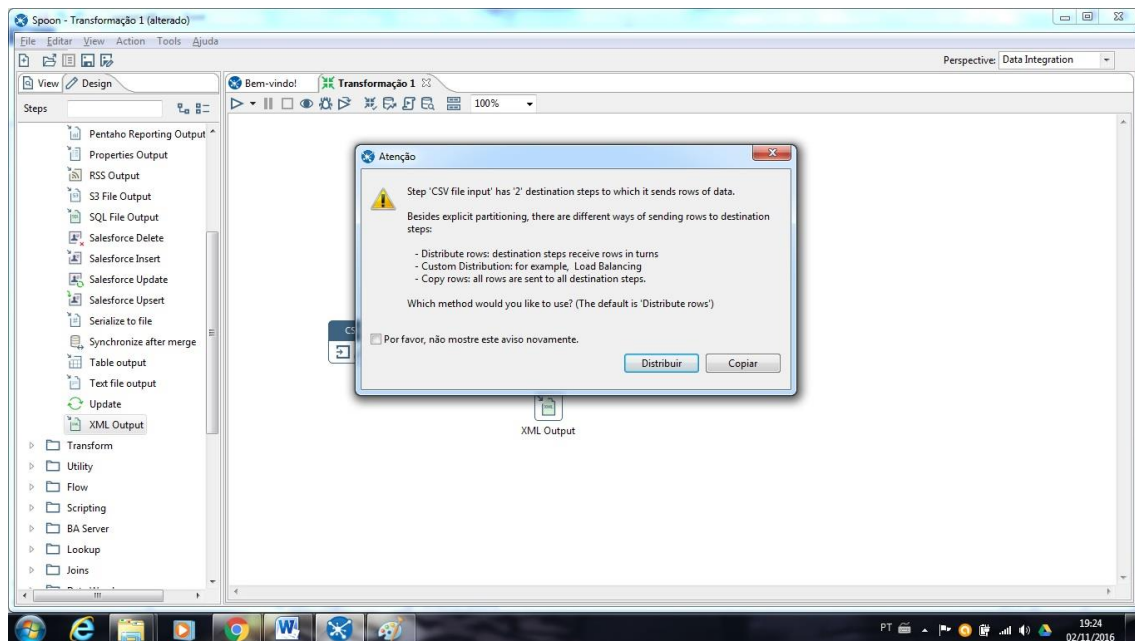
Em seguida, criaremos as conexões. Clicando em cada saída e escolhendo a primeira opção: . Em seguida arrastamos a seta exibida até o ícone de Input.



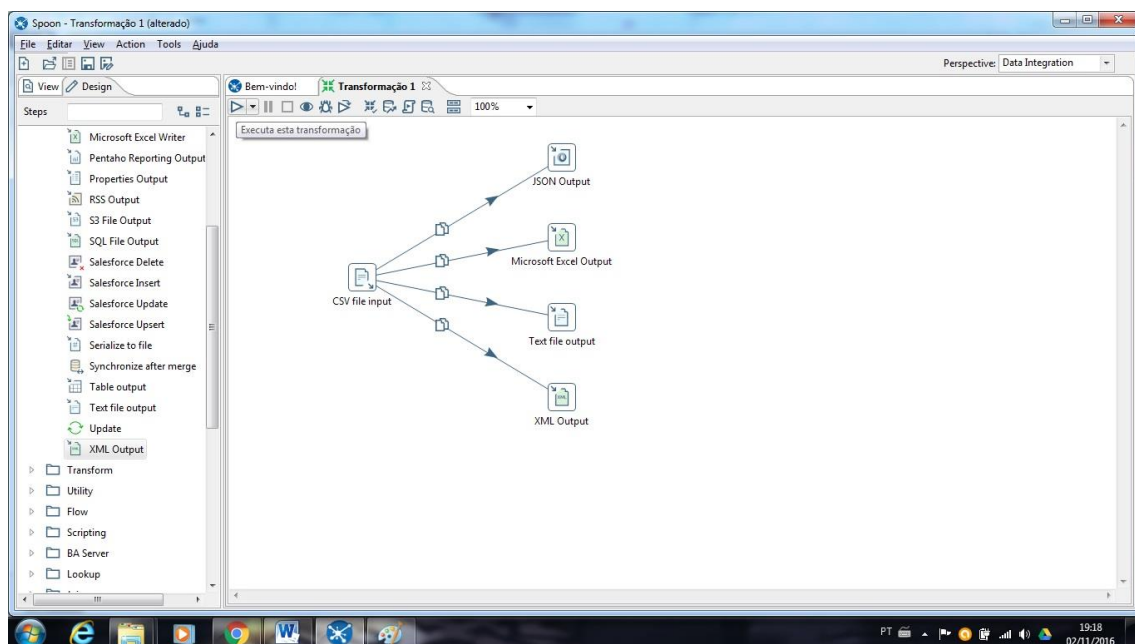
Cada conexão, chamamos de Hop. A cada Hop criado, escolhemos a opção: , na janela exibida abaixo.



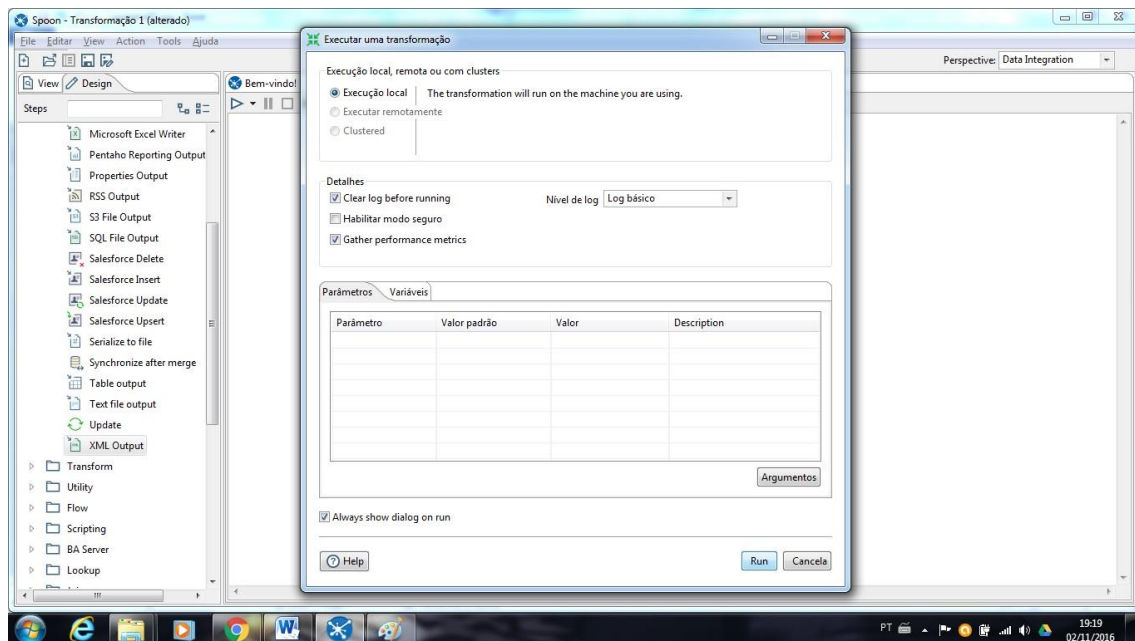
Como, escolhemos 4 opções de saída, o PDI exibirá a opção de gerar cópias, basta clicar em “Copiar”, como visto abaixo.



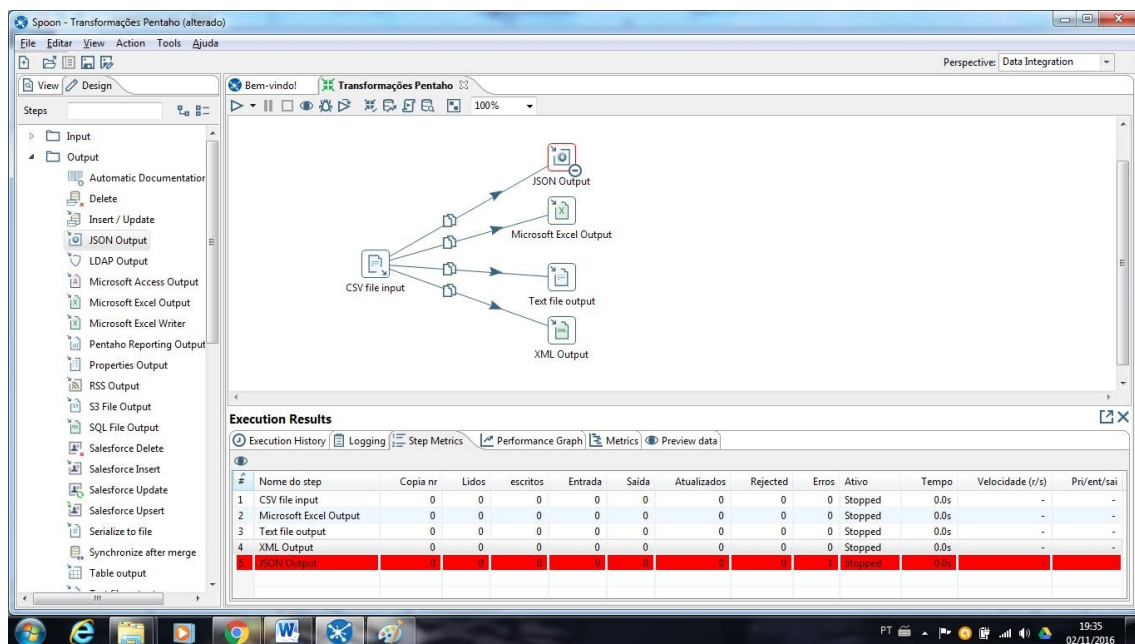
O próximo passo é executar transformação. Clicando no ícone descrito na imagem:



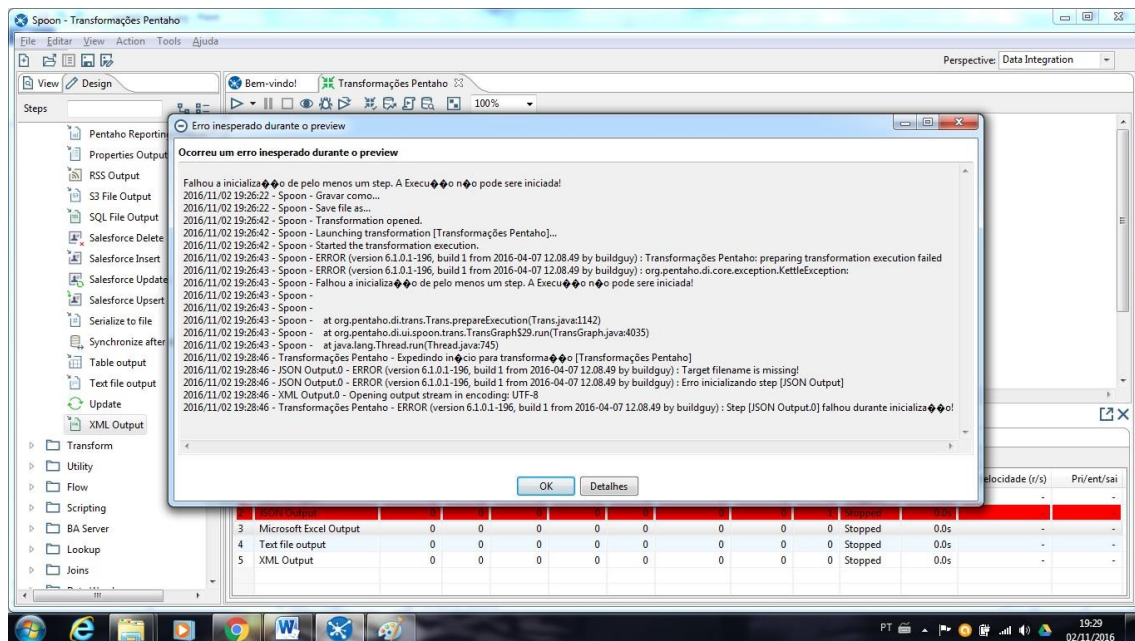
Em seguida clicamos em “Run”, como visto na janela a seguir.



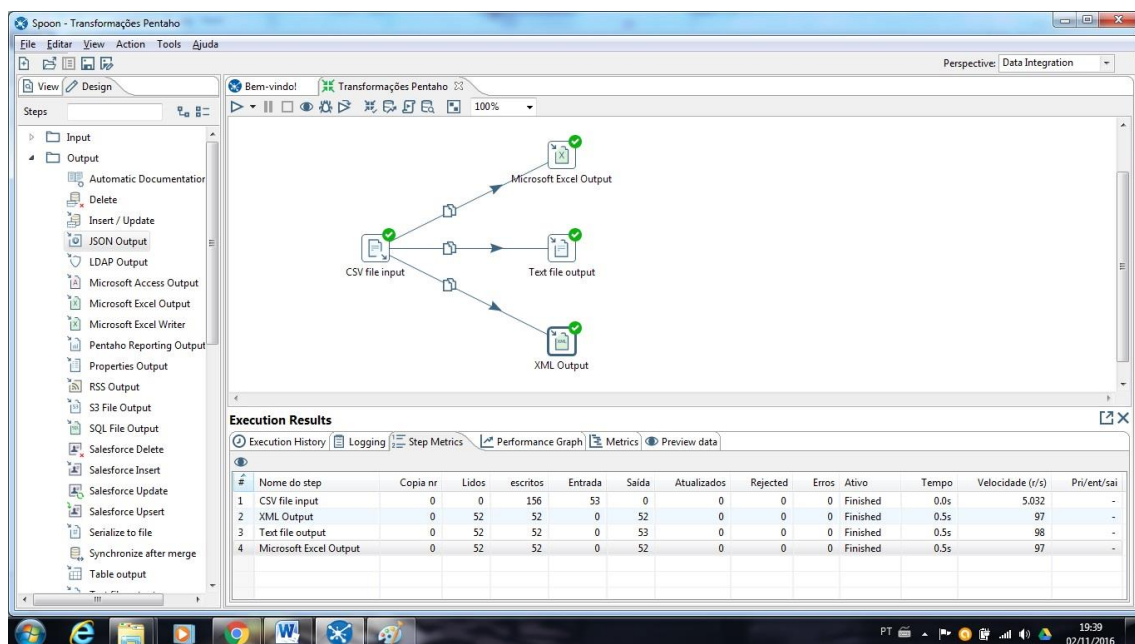
Um erro inesperado foi encontrado na transformação para o formato JSON.







Após a remoção dessa saída, refazendo os passos anteriores, a transformação foi realizada com sucesso.



O Arquivo de transformação foi gerado com sucesso.

