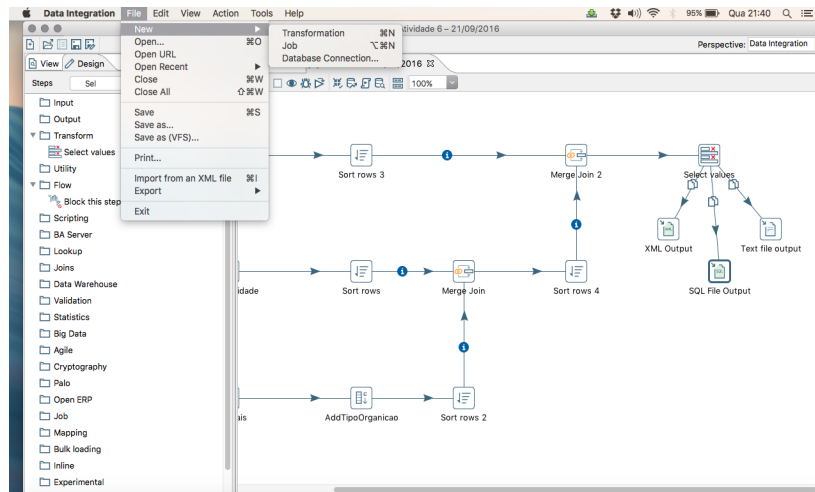


**UNIVERSIDADE FEDERAL DE SERGIPE**  
**DEPARTAMENTO DE COMPUTAÇÃO**  
**SISTEMAS DE APOIO À DECISÃO**  
**ANNE CAROLINE MELO SANTOS**

**Tutorial Atividade 7**

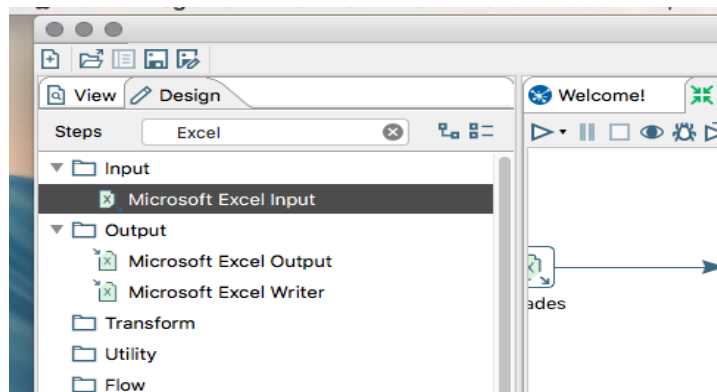
1º Passo: Abra a aplicação PDI

2º Passo: Crie uma nova transformação clicando nas seguintes opções em sequência **File>New>Transformation**



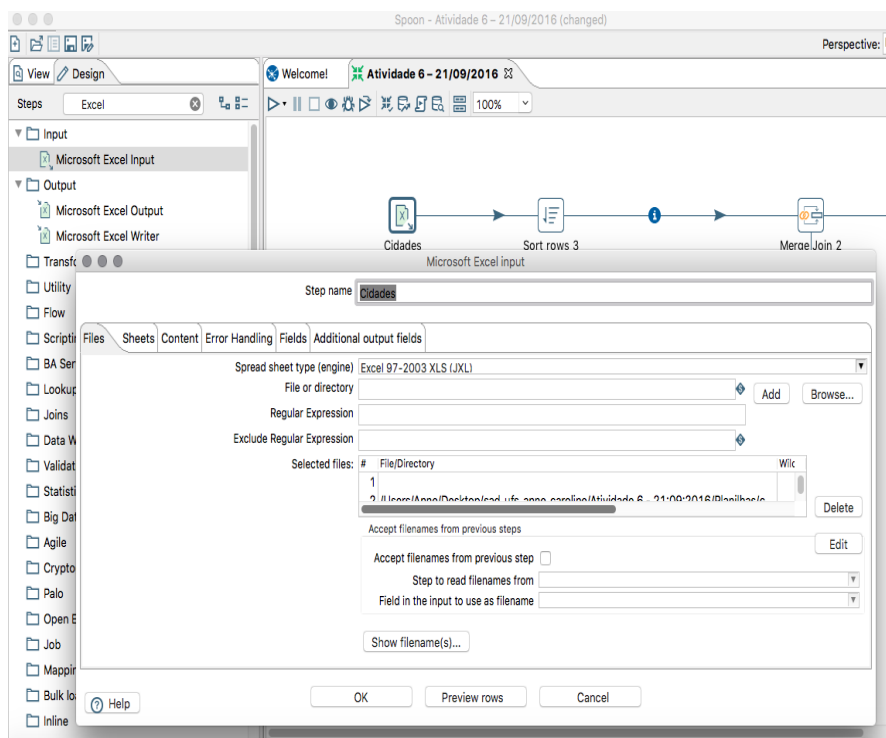
**Figura 1:** Criando uma Transformação.

3º Passo: Clique na aba **Design** e abra categoria **Input**. Na categoria **Input**, clique na opção **Microsoft Excel Input** e arraste-a até a tela de transformação.



**Figura 2:** Criando Inputs Excel.

4º Passo: Defina um nome para o **Input** que está recebendo o arquivo excel. Atribua ao **Input** o arquivo excel que está alocado em sua máquina, seguindo os seguintes procedimentos:  
**Browser>Open>Add**



**Figura 3:** Configurando o Input de entrada.

5º Passo: Ainda nas configurações do **Microsoft Excel Input**, clique na aba **Sheets** e depois em **Get sheetnames**.

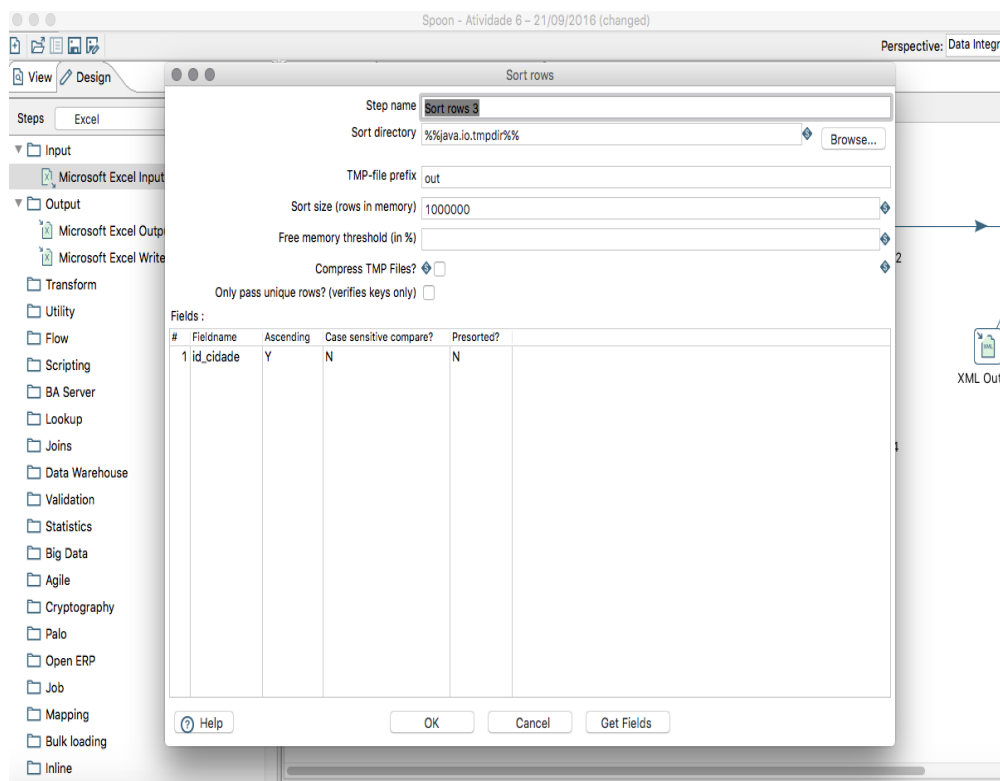
6º Passo: Defina os campos que serão importados da planilha de entrada, clicando na aba **Fields** e depois clicando na opção **Get fields from header row**. Após a importação dos campos, altere o **type** dos campos para tipos que sejam condizentes com sua necessidade. Por exemplo, em vez da coluna id ser **Number**, substitua por **Integer**.

7º Passo: Clique em **Preview rows** para verificar se a importação foi feita corretamente.

8º Passo: Caso os dados apareçam corretamente, clique em **close** para fechar a pré-visualização e depois clique em **OK** para concluir a configuração do seu **Input**.

9º Passo: Aplique os passos do 3º ao 8º mais duas vezes, pois estamos lidando com três entradas excel diferentes.

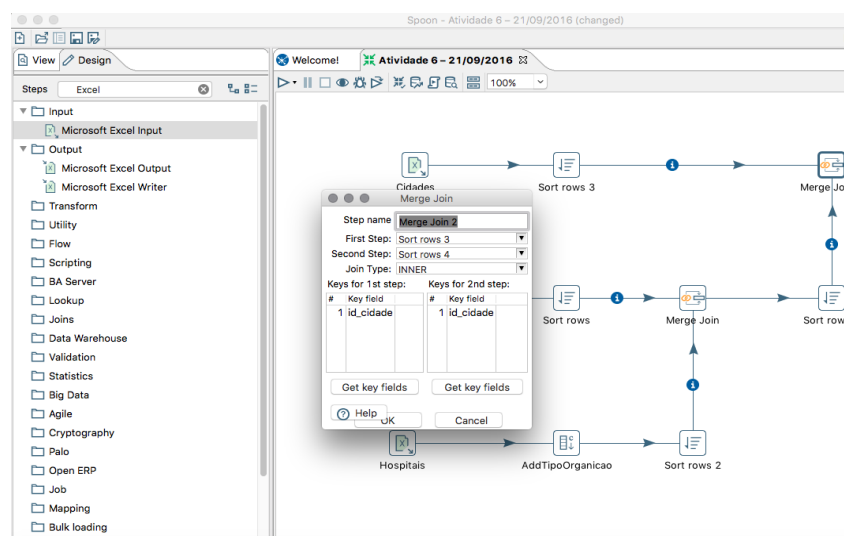
10º Passo: Para fazer o **merge** entre as planilhas de entrada, é necessário que os ids estejam em ordem ascendente. Para tal siga os seguintes passos: na categoria **Transform**, escolha a opção **Sort rows** e arraste-a para tela de transformação. Ligue o **Microsoft Input Excel** ao **Sort rows**. Clique **Sort rows** e abra a caixinha diálogo correspondente a este componente. Em **Fieldname**, selecione o campo que deve estar em ordem crescente, na aba **Ascending** escolha a opção Y. Feito isso clique em **OK**.



**Figura 4:** Colocando em ordem crescente.

11º Passo: Repita o décimo passo para os outros inputs da transformação, até que todas as planilhas tenham seus ids ordenados.

12º Passo: Após a ordenação das planilhas de entrada, na aba **design**, na categoria **Calculator**, escolha a opção **merge** e arraste-a para tela de transformação, ligue o **Sort rows** ao **merge**. Na caixa de diálogo que abre, defina o nome do step, defina o **First Step** e **Second Step**, em **Join Type** escolha a opção **INNER**. Defina a chave de interseção entre as duas planilhas, depois clique em **OK**.



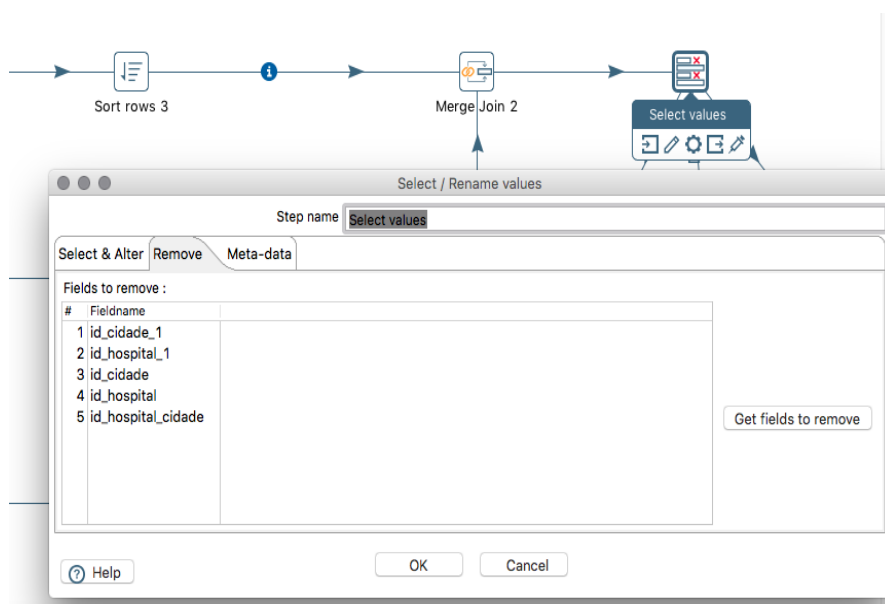
**Figura 5:** Configurando o merge.

13º Passo: Clique no ícone do olho para ter uma pré-visualização do seu **merge** e conferir se ele está correto.

14º Passo: Repita o 14º passo, até conseguir a interseção entre todas as planilhas e assim chegar ao resultado desejado.

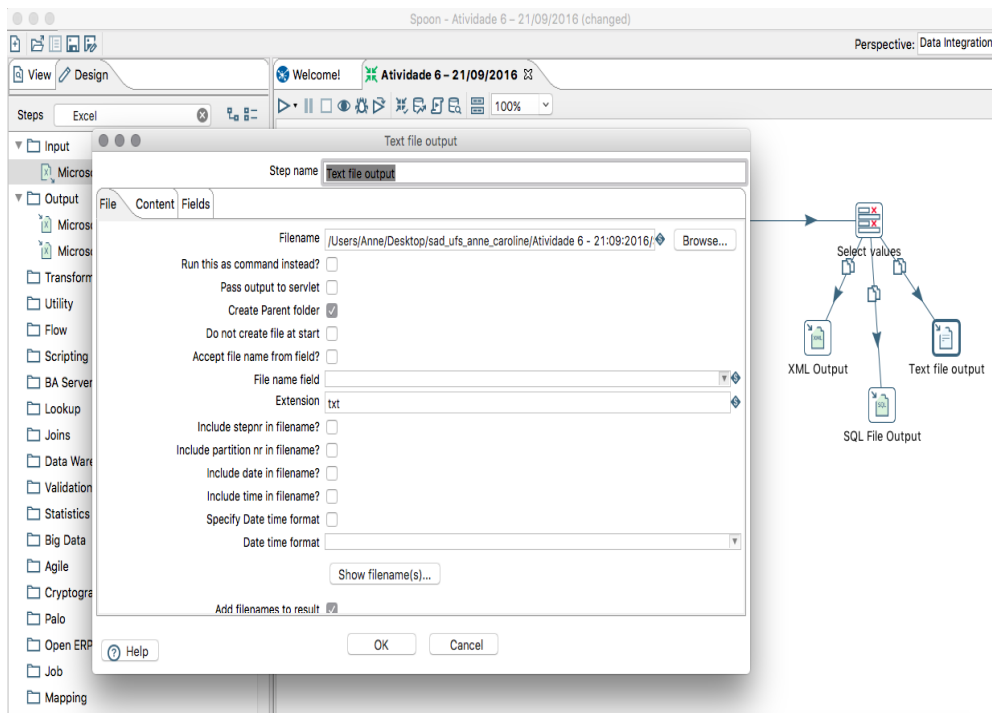
15º Passo: Realizado o **merge**, é importante verificar se o resultado está ordenado, caso não esteja, reaplique o 10º passo para poder ordenar a lista.

16º Passo: Após a ordenação do **merge**, na aba **Transform**, escolha a opção **Select Values** e arraste-a para tela de transformação. Feito isso, arraste uma seta do **merge** para o **Select values**. Clique duas vezes no **Select Values**, na caixa de diálogo que abre defina um nome para o arquivo de saída e defina quais são os campos que não devem permanecer, mantendo apenas as informações que lhe são úteis.



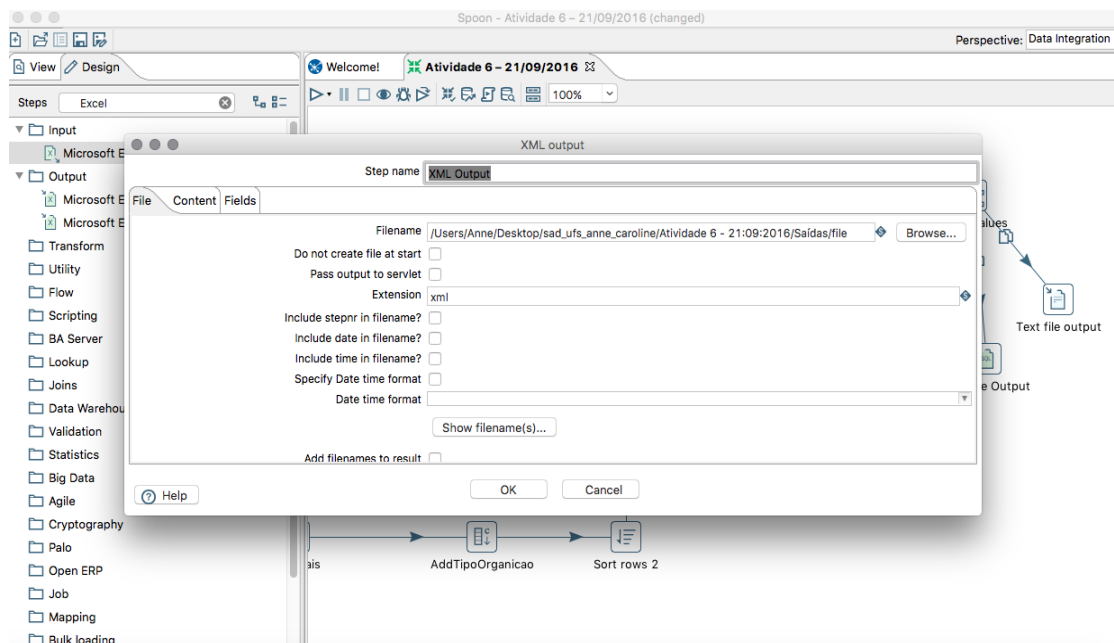
**Figura 6:** Configurando o select values.

17º Passo: Após definir os campos que devem ser removidos, é importante gerar as saídas. Trabalharemos com três tipos de saída: **txt**, **xml** e **sql**. Na aba **output**, escolha a opção **Text file output** e arraste-a para tela de transformação. Feito isso, arraste uma seta do **Select values** para o **Text file output**, clique duas vezes sobre o **output txt**, na tela que abre defina o nome para o arquivo que será gerado e quais serão os campos que estarão presentes no documento, clicando em **Get Fields**. Concluída o preenchimento, clique em **OK**..



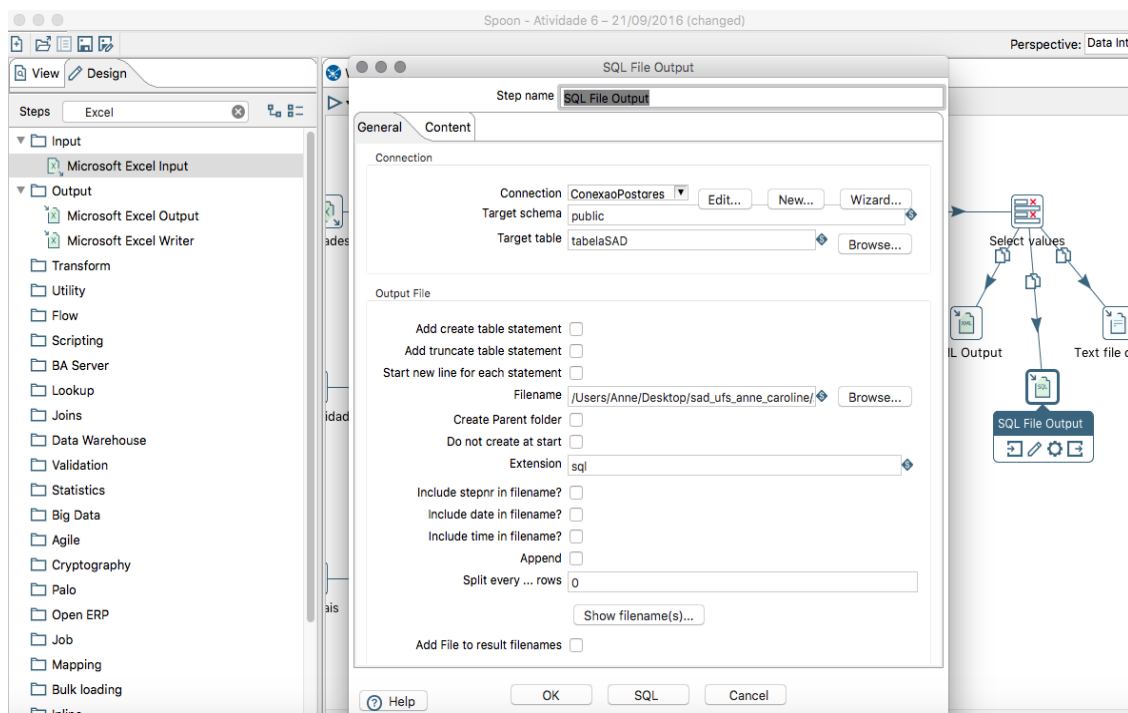
**Figura 7:** Configurando a saída .txt

18º Passo: Gerando uma saída xml. Na aba **output**, escolha a opção **XML output** e arraste-a para tela de transformação. Feito isso, arraste uma seta do **Select values** para o **XML output**, clique duas vezes sobre o **output xml**, na tela que abre, defina o nome para o arquivo que será gerado e quais serão os campos que estarão presentes no documento, clicando em **Get Fields**. Concluído o preenchimento, clique em ok.



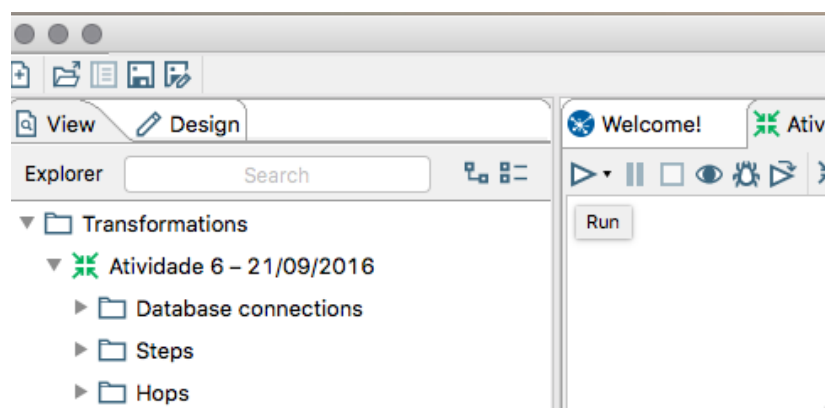
**Figura 8:** Configurando a saída .xml

19º Passo: Gerando uma saída Banco de Dados. Na aba output, escolha a opção **SQL file output** e arraste-a para tela de transformação. Feito isso, arraste uma seta do **Select values** para o **SQL file output**, clique duas vezes sobre o **output sql**, na tela que abre, defina o nome para o arquivo que será gerado, e uma nova conexão clicando em new. Preencha os dados de conexão de acordo com suas preferências, defina qual o nome do schema, e o nome da tabela, onde os dados serão inseridos. Clique em SQL para que a query seja executada, depois clique em **OK** para que o script seja gerado e alocado em uma pasta em seu computador.



**Figura 9:** Configurando saída sql.

20º Passo: Geradas as três saídas, clique em **Run** e veja se a transformação aconteceu corretamente.



**Figura 10:** Executando a transformação.

21º Passo: Caso tudo tenha funcionado, parabéns, você acabou de realizar um procedimento **ETL**.