

TUTORIAL PENTAHO LIPPAUS

Sumário

1.	Instalação Pentaho.....	4
1°	Passo – Baixar Pentaho.....	4
2°	Passo – Cadastrar e fazer login no Oracle:	4
3°	Passo – Baixar JDK8	4
4°	Passo – Cadastrar Variável Pentaho	4
2.	Configurar Pentaho	4
1°	Passo – Configurar Variáveis de Sistema	4
3.	Criar Ambiente de Trabalho	5
1°	Passo – Criar Transformação	5
	- Vai em “File” -> “Transformações” OU (Ctrl + N).....	5
2°	Passo – Criar Conexões (DBMaker)	5
	- Vai em “View” -> “Conexões” -> “Novo”	5
2.1°	Passo – Criar Conexões (output)	5
	- Vai em “View” -> “Conexões” -> “Novo”	5
1°	DICA: Como Salvar.....	6
4.	Criar Select	6
1°	Passo - Criar Select do Input	6
2°	Passo - Criar Select do output	6
5.	Criar Filtro de Dados (por coluna)	6
1°	Passo – Aplicar na Área de Trabalho	6
2°	Passo – Configurar o Select Values.....	7
6.	Teste de Dados	7
1°	Passo – Testar os Dados	7
7.	Como Lançar Dados Para o Bando (output).....	7
2°	DICA: Como Limitar Quantidade de Linhas de um Campo	7
3°	DICA: Como dar Comando Para um Banco de Dados	7
1°	Passo – Criar Script	7
2°	Passo – Configurar Script.....	7
8.	Como utilizar 1 Select para mais de 1 caminho	8
9.	Como Criar e Utilizar um Join.....	8

1° Passo – Adicionar Join na área de trabalho.....	8
2° Passo – Como Configurar o Join	8
-	8
4° Dica: Em caso de Dados Muito Grande (Onde o pentaho se confunde)	8
10. Como Exportar Dados para Formato EXCEL	9
1° Passo – Adicionar a área de trabalho	9
2° Passo – Configurar Exportação Excel.....	9
2.1° Passo – Obter Colunas Excel.....	9
10.1. Como Importar Dados do Excel para o Banco.....	9
1° Passo – Adicionar a Área de Trabalho	9
2° Passo – Configurar o Microsoft Excel Input.....	9
2.1° Passo – Configurando os Dados que Serão Enviados.....	9
10.2. Como Tratar Dados Importados do Excel.....	10
11. Como Fazer Cópia de Uma Área de Trabalho	10
12. Como Ler o Banco de Dados em JSON	10
1° Passo – Colocando arquivos na área de trabalho	10
2° Passo – Configurar JSON.....	10
3° Passo – Gerar JSON.....	11
12.2. Ler JSON.....	11
1° Passo – Colocando arquivos na Área de Trab.	11
- 3° Passo – Importar JSON	11
13. Criação Job e Configuração	11
1° Passo – Criar Job	11
2° Passo - Inserindo Arquivos	12
3° Passo – Configurando email	12

1. Instalação Pentaho

1º Passo – Baixar Pentaho

- Entre no Link: <https://sourceforge.net/projects/pentaho/>
- Click em download

2º Passo – Cadastrar e fazer login no Oracle:

- Entre no link: <https://profile.oracle.com/myprofile/account/create-account.jspx>
- Criar conta no site Oracle.
- Inicie a sessão: <https://login.oracle.com/mysso/signon.jsp>

3º Passo – Baixar JDK8

- Entre no Link: <https://www.oracle.com/br/java/technologies/javase/javase-jdk8-downloads.html>
- Procure o compatível com Windows X64
- Baixe o programa

4º Passo – Cadastrar Variável Pentaho

- Click a tecla **“Windows”**
- Pesquise: **“Variável”** -> **“Editar as variáveis de am...”** -> **“Variáveis de ambiente”** -> **“PENTAHO...”**
- Mude o caminho para onde esta o JDK8

2. Configurar Pentaho

1º Passo – Configurar Variáveis de Sistema

- Click Junto as teclas: (**Ctrl + Alt + P**)
- Click com o botão direito no primeiro numero -> **“Insert before th...”**
- De um nome uma (**Palavra/Valor**) equivalente

3. Criar Ambiente de Trabalho

1º Passo – Criar Transformação

- Vai em “File” -> “Transformações” OU (Ctrl + N)

2º Passo – Criar Conexões (DBMaker)

- Vai em “View” -> “Conexões” -> “Novo”

Insira os dados abaixo:

Connection name = **“DBMaker1”**

Connection type = **“Generic database”**

Access = **“Native (JDBC)”**

Settings:

Dialect = **“Generic data...”**

Custom connection URL = **a.db_url**

Custom drive... = **a.db_drive**

Authentication:

Username: **a.db_user**

Senha: **a.db_senha**

2.1º Passo – Criar Conexões (output)

- Vai em “View” -> “Conexões” -> “Novo”

Insira os dados abaixo:

Connection name = **“DW”**

Connection type = **“PostgreSQL”**

Access = **“Native (JDBC)”**

Settings:

Host name = **a.dw_ip**

Database Name = **a.dw_database**

Port Number = 5432

Username: **a.dw_user**

Password: **a.dw_senha**

- **“OK”**

3° Passo – Gerar Table

- Design -> Pesquise **“table”**
- Arraste: **“Table Input”** para a área de trabalho do Pentaho
- Faça o mesmo com **“Table Output”**

1° DICA: Como Salvar

- Click junto as teclas: (**Ctrl + S**)
- Escolher local e criar nova pasta
- Salvar dentro da pasta

4. Criar Select

1° Passo - Criar Select do **Input**

- Click 2 vezes no **“Table Input”**
- Mude a conexão para **“DBMaker”**
- Crie o select no **“SQL”**

2° Passo - Criar Select do **output**

- Click 2 vezes no **“Table output”**
- Mude o nome em **“Target table”**
- Marque **“User batch update...”** caso **NÃO** esteja marcada

(**OBS: “SQL” SERVE PARA VISUALIZAR OS DADOS QUE ESTÃO SAINDO**)

5. Criar Filtro de Dados (por coluna)

1° Passo – Aplicar na Área de Trabalho

- Vai em **“Design”** e arraste **“Select Values”** para a área de trabalho em cima da linha que vincula o **Input** e **Output**

2º Passo – Configurar o Select Values


- Click 2 vezes em **“Select Values”** -> **“Get Fields to Select”**
- Mude o tipo do campo em **“Meta-data”** -> **“Format”** e selecione o tipo desejado

(OBS: Pode renomear os dados em **“Rename to”**)


6. Teste de Dados

(OBS: ANTES DESVINCULE A LINHA QUE LIGA A ETAPA FINAL (EXEMPLO: OUTPUT))

1º Passo – Testar os Dados

- Click no botão []
- Click em **“Run”**

7. Como Lançar Dados Para o Bando (output)

- Verifique se todas os vínculos estão funcionando (**NÃO** pode estar transparente)
- Click no botão []
- Click em **“Run”**

2º DICA: Como Limitar Quantidade de Linhas de um Campo

- Click 2 vezes no **“Input”** que deseja limitar
- Adicione no **“SQL”** -> **Limit (quantidade de linhas)**

3º DICA: Como dar Comando Para um Banco de Dados

1º Passo – Criar Script

- Vai em **Design** -> Pesquise: **“Scripting”** e arraste **“Execute SQL Script”**

2º Passo – Configurar Script

- Click 2 vezes no **“script”**
- Em **“Connection”** escolha o banco de dados que deseja dar o comando
- Escreva o comando

EXEMPLO: Vamos dar um comando para deletar os dados **antigos** contidos no banco

Utilize o comando: Delete from (Nome do campo desejado)

(OBS: “ ; ” NÃO é obrigatório em caso de 1 script)

8. Como utilizar 1 Select para mais de 1 caminho

- Vincule o select com outro desejado em vez de clicar em “**Distribuir**” click em “**Copiar**”

O QUE “DISTRIBUIR” FAZ??? Divide os dados informados entre os vínculos

O QUE “COPIAR” FAZ??? Envia todos os dados igualmente para todos os vínculos

9. Como Criar e Utilizar um Join

1º Passo – Adicionar Join na área de trabalho

- Vá em Design -> Pesquise: “**Join**” e arraste “**Merge join**” para a área de trabalho em frente aos **Selects** que deseja juntar os dados

2º Passo – Como Configurar o Join

- Vincule os **Selects** ao “**Join**”

“**Fist Step**”: qual dado será o primeiro.

“**Second Step**”: qual dado será o segundo.

[E assim por diante...]

4º Dica: Em caso de Dados Muito Grande (Onde o pentaho se confunde)

- Adicione um “**sort rows**” na área de trabalho para cada **select Input**.

- Click 2 vezes no **Step** e em “**Obter campos**”

- Selecione os dados que **VAI** passar. (Deletando os indesejados)

10. Como Exportar Dados para Formato EXCEL

1º Passo – Adicionar a área de trabalho

- Adicione a área de trabalho **“Microsoft Excel Output”**
- Vincule o input com os dados desejados ao **“Microsoft...”**

2º Passo – Configurar Exportação Excel

- Click 2 vezes no **“Microsoft...”**
- Selecione o local de destino em **“Navegar...”**
- Modifique para o nome desejado em **“File”**

2.1º Passo – Obter Colunas Excel

- Vá em **“Field”** -> **“Obter Campos”**.
- Mude o tipo de campo que vai sair e **“Type”** (Caso necessário).
- **“OK”**

10.1. Como Importar Dados do Excel para o Banco

1º Passo – Adicionar a Área de Trabalho

- Arraste o **“Microsoft Excel Input”** para a área de trabalho
- Vincule ao **output** com o banco de dados que deseja inserir os dados do **Excel**.

2º Passo – Configurar o Microsoft Excel Input

- Click 2 vezes no **“Microsoft...”**
- Selecione o local do arquivo que deseja importar em **“Navegar...”**
- Selecione a versão do seu Excel em **“File or directory”**

2.1º Passo – Configurando os Dados que Serão Enviados

- Vá em **“Sheets”** -> **“Get Sheetname(S)...”**
- Selecione suas **Sheets** e utilize o botão **“>”** para selecionar quais **Sheets** serão enviadas
- Vá em **“Fields”** e selecione os campos que serão enviados.
- **“OK”**

(OBS: Para verificar se os dados desejados estão sendo enviados)

- Click em "Microsoft..."
- Click em "Preview..."
- "OK"

10.2. Como Tratar Dados Importados do Excel

(OBS: Podemos tratar os dados utilizando "strings" ou "java")

EXEMPLO: Vou Tratar o dados Utilizando "String" (Uma das Formas)

- Arraste "Replace in String" para a área de trabalho (a frente do "Microsoft...")

(OBS: Mude o tipo do dado desejado utilizando "Select Values" para string (Caso NÃO seja uma String) pois para tratar os dados desta forma, precisa ser String)

- Click 2 vezes em "Replace in String"
- Em "Seach" Buscamos o dado que deseja mudar
- Em "Replace with" será o dado que vai virar.
- "OK"

11. Como Fazer Cópia de Uma Área De Trabalho

- Vai em "Salvar como"
- Salve na mesma pasta onde se encontra as outras áreas de trabalho

(OBS: Assim todas as conexões a acompanharam)

12. Como Ler o Banco de Dado em JSON

1º Passo – Colocando arquivos na área de trabalho


- Arraste "JSON Output" para a área de trabalho
- Faça o mesmo com "Table Input"

2º Passo – Configurar JSON

- Click vezes em "JSON..."
- Compatibility mode: Desative (caso esteja marcado)
- Mude o destino do arquivo em "Navegar..."
- Mude a "extensão" para "From"

- Fields -> **“Get ...”** -> Delete o campos indesejados.
- **N... rows in a bloc** = 0

3º Passo – Gerar JSON

- Click no botão []
- Click em **“Run”**

12.2. Ler JSON


1º Passo – Colocando arquivos na Area de Trab.

- Arraste **“JSON Input”** para a área de trabalho.
- Arraste **“Table Output”** para a área...

2º Passo – Configurando Arquivos

- Click 2 vezes no **“JSON...”**
- **Browse** -> Escolha eu arquivo **“JSON”**
- **Add**
- **Fields** -> **“Select Fields”** -> Selecione os campos desejados -> **OK**
- **“OK”**

- 3º Passo – Importar JSON

- Click no botão []
- Click em **“Run”**

13. Criação Job e Configuração

1º Passo – Criar Job

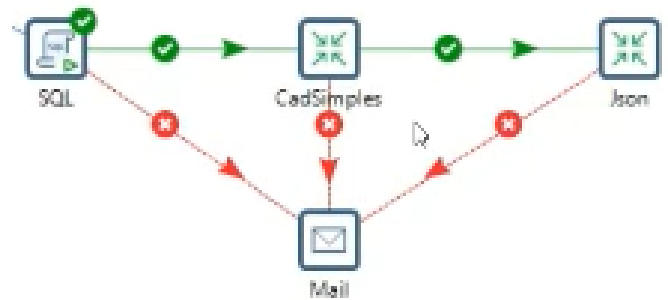
- Click em **“Novo”**
- Click em **“Job”**

2º Passo - Inserindo Arquivos

- Arraste **"Start"** para a área de trab.
- Arraste **"Transformação" (quantas precisar)**
- Arraste **"SQL"** para a área de trab.
- Arraste **"Mail"** para a are...
- Vincule: Start -> SQL -> Trab1... ->
- Modifique o tipos de vínculos clicando nos ícones :

Sendo:  Se **NÃO** haver erros

 Se **HAYER** erros



3º Passo – Configurando email

- Click 2 vezes no **"Mail"**
- **Destination Address** = [email desejado]
- Sender name = [Email detinado desejado]
- ender addres = []

Servers

SMTP Server: Smtg.gmail.com

Porta: [Verificar pesquisando no navegador = **SMTP**]

Authentication user : [Email desejado]

Authentication password : [Senha]

User secure Authentication: Marcado

- Vai em **"Attached Files"**: Marque

Message:

Suject = Erro

Comment = eternal. Job . file fllename Name

4º Passo – Configurando transformações

- Click 2 vezes na transformação desejada
- "Browse..." [Transformação desejada] (faça isso em ordem da sequência)

5º Passo – Agendar Job

- Vamos na pasta onde foi salvo nossos trabalhos
- Edite o **"Start"**
- Mude o caminho para o **"job"** dentro da pasta do nossos trabalho
- Reflete o mesmo com o **"Log"**