

Plataforma Pentaho de Business Intelligence

Manual de utilização

Universidade Federal da Bahia
Departamento de Ciência da Computação

Sumário

1. Pentaho BI.....	3
2. Instalação e Utilização.....	4
1. Pré-requisitos	4
2. Instalação do Pentaho BI platform	4
3. Área Administrativa	5
1. Criar nova Conexão de dados	5
3. Definição da arquitetura	6
1. Construção de modelo Multidimensional	6
1. Instalação e Configuração	6
2. Criação do modelo e geração da base de dados	7
2. Criar o schema para identificação do cubo OLAP	8
1. Schema Workbench	8
1. Funcionalidades	8
2. Instalação e Configuração	8
3. Criando uma conexão com o Banco de dados.....	8
4. Criando um Schema para identificação de um cubo	8
2. Exportando o Schema criado para o Pentaho	10
4. Processo de ETL (Extraction, Transformation and Load)	12
1. Download e Instalação	13
2. Configuração	13
3. Utilização do PDI	15
1. Conceitos importantes.....	15
5. Análise do Cubo	16
1. Abrindo Schema criado anteriormente	16
6. Adicionando Usuários	17
1. Editando Usuários.....	17
2. Deletando Usuários.....	17
3. Adicionando e excluindo Papéis	17
7. Personalização do Pentaho	18
1. Tradução do Pentaho	18
8. Outras ferramentas.....	18
1. Pentaho Analysis Tool (PAT).....	18
2. Pentaho Analyser	19
3. Pentaho Report Designer (PRD).....	20
4. Web Ad Hoc Query and Reporting Service (WAQR).....	21
5. Community Dashboard Framework	22
9. Créditos	23
10. Referências	24

Pentaho BI

A plataforma Pentaho, conhecida como Pentaho Open BI, é uma suite aplicativos Open Source para criação de soluções de BI (Business Intelligence). Nesse tutorial será detalhado a instalação e utilização da versão 3.5 stable do Pentaho.

Os módulos mais importantes do Pentaho são:

- Pentaho BI Platform;
- Pentaho Data Integration - PDI (Kettle);
- Analysis View (Mondrian);
- Pentaho Reporting (Reporting) e;
- Weka (Data Mining).

Nesta versão (3.5) o Pentaho BI Platform, o Analysis View e o Pentaho Reporting vêm integrados no mesmo pacote "bi-server". Para adquirir as outras ferramentas é necessário o download individual e a configuração no ambiente. É importante ressaltar que o Pentaho pode ser utilizado no Windows, Linux, Mac, BSD, Solaris, etc.

Além dessas ferramentas listadas, existem outras que são conhecidas pela comunidade do Pentaho e aos poucos estão sendo agregadas.

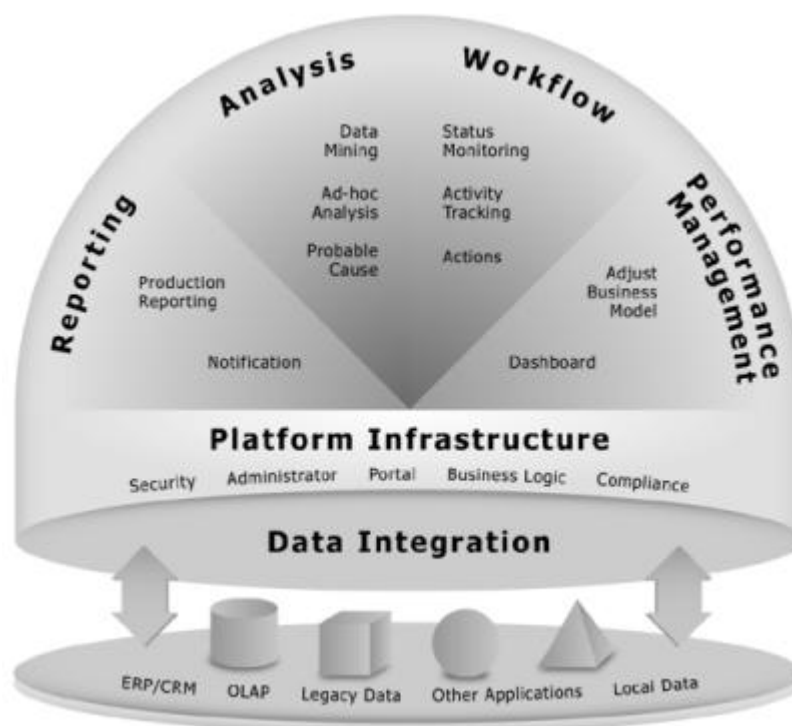


Figura 1 - A arquitetura do Pentaho

A plataforma Pentaho BI é distribuída através da Licença Pública da Pentaho (PPL - Pentaho Public License) de código aberto. Maiores informações sobre podem ser encontradas em <http://www.pentaho.com/license/>.

Existem duas formas disponíveis para adquirir a ferramenta. A versão Community pode ser baixada gratuitamente em <http://sourceforge.net/projects/pentaho/files/>. Só nesse ano foram lançadas ou sofreram atualização de código pelo menos quatro versões, o que mostra o grande desenvolvimento da comunidade livre do Pentaho. Além dessa, existe a versão Enterprise. O download de uma versão trial pode ser efetuado no

próprio site do Pentaho (<http://www.pentaho.com/>) para testes durante 30 dias. A grande diferença entre as duas versões, além do preço, é que a versão Enterprise possui uma suíte de ferramentas já previamente integradas e algumas mais desenvolvidas que a versão livre. Além disso, contém suporte técnico, manutenção com atualizações frequentes, acesso a melhor documentação e treinamento. Um comparativo pode ser visto em http://www.pentaho.com/products/enterprise/enterprise_comparison.php. Em relação aos valores, no período que foi feito esse tutorial, o valor para adquirir a suíte integrada do Pentaho Enterprise era de \$30.000 USD/EUR.

Instalação e Utilização

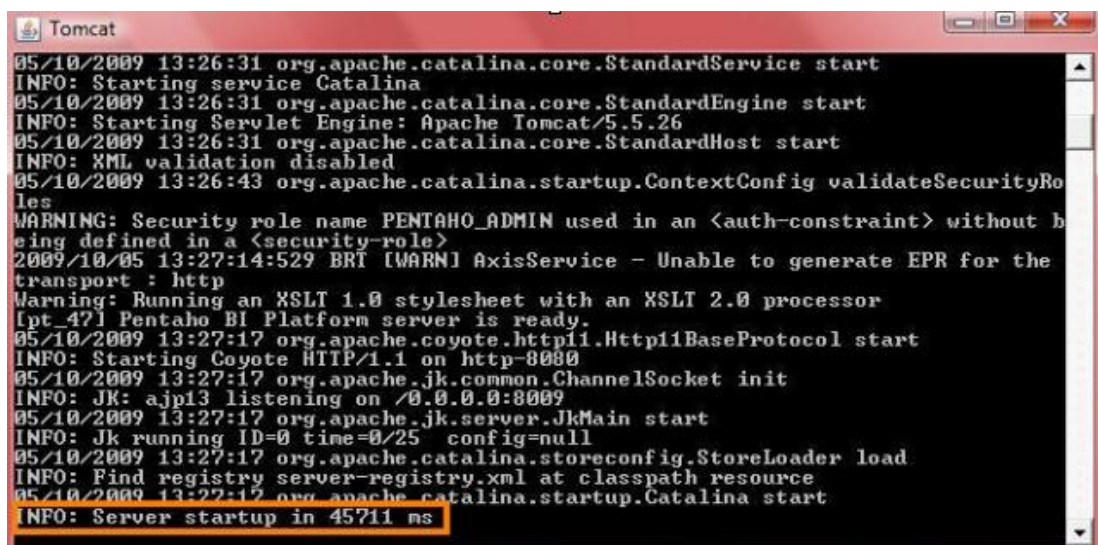
Neste tutorial vamos mostrar a instalação do Pentaho e seu uso básico. Para isso vamos criar um modelo de dados dimensional em um banco de dados PostgreSQL 8.3. Vamos também simular o ETL a partir de um banco de dados MySQL, de onde estaria nosso modelo Entidade Relacionamento (base transacional).

Pré-requisitos

Ter o ambiente Java instalado e configurado. Definir as variáveis de ambiente JAVA_HOME ou JRE_HOME.

Instalação do Pentaho BI platform

1. Baixe o arquivo [biserver-ce-3.5.0.stable.zip](#) (Windows) ou [biserver-ce-3.5.0.stable.tar.gz](#) (Linux, Mac, etc).
2. Descompacte o arquivo em uma pasta de sua preferência.
3. Para iniciar o Pentaho execute o arquivo chamado start-pentaho.bat (Windows) ou start-pentaho.sh (Linux) na pasta "biserver-ce/" e aguarde a inicialização. O pentaho é executado sobre um Servidor Web, que por padrão é o tomcat. Deve aparecer a mensagem "INFO: Server startup in <tempo> ms" quando finalizar a inicialização.



```
05/10/2009 13:26:31 org.apache.catalina.core.StandardService start
INFO: Starting service Catalina
05/10/2009 13:26:31 org.apache.catalina.core.StandardEngine start
INFO: Starting Servlet Engine: Apache Tomcat/5.5.26
05/10/2009 13:26:31 org.apache.catalina.core.StandardHost start
INFO: XML validation disabled
05/10/2009 13:26:43 org.apache.catalina.startup.ContextConfig validateSecurityRoles
WARNING: Security role name PENTAH0_ADMIN used in an <auth-constraint> without being defined in a <security-role>
2009/10/05 13:27:14:529 BR1 [WARN] AxisService - Unable to generate EPR for the transport : http
Warning: Running an XSLT 1.0 stylesheet with an XSLT 2.0 processor
[pt_47] Pentaho BI Platform server is ready.
05/10/2009 13:27:17 org.apache.coyote.http11.Http11BaseProtocol start
INFO: Starting Coyote HTTP/1.1 on http-8080
05/10/2009 13:27:17 org.apache.jk.common.ChannelSocket init
INFO: JK: ajp13 listening on /0.0.0.0:8009
05/10/2009 13:27:17 org.apache.jk.server.JkMain start
INFO: Jk running ID=0 time=0/25 config=null
05/10/2009 13:27:17 org.apache.catalina.storeconfig.StoreLoader load
INFO: Find registry server-registry.xml at classpath resource
05/10/2009 13:27:17 org.apache.catalina.startup.Catalina start
INFO: Server startup in 45711 ms
```

Figura 2 - Console durante execução do Pentaho

4. Acesse o endereço <http://localhost:8080/pentaho> no seu navegador. Clique no botão Login, ele oferece alguns usuários pré-configurados. O administrador é o usuário "Joe".
5. O Pentaho já vem com uma base de dados de testes pré-configurada e alguns exemplos de relatórios e views criadas. É possível conhecer um pouco da ferramenta navegando pelo menu na lateral esquerda.

Obs.: você pode observar o log do Tomcat no arquivo "biserver-ce/tomcat/logs/catalina.out" para identificar possíveis erros durante as etapas seguintes.


Área Administrativa

O Pentaho possui um painel administrativo em que é possível gerenciar contas de usuários e conexões com banco de dados. Para acessar a área administrativa devemos:

- Iniciar o serviço em: start-pac.sh (Linux) ou start-pac.bat (Windows), na pasta "administration-console/"
- Para parar o serviço: stop-pac.sh ou stop-pac.bat.
- Quando o console exibir a mensagem "INFO: Console is now started. It can be accessed using ...", a página do administrador poderá ser acessada no endereço: <http://localhost:8099/>
- O nome do usuário padrão é "admin" e a senha "password".

Criar nova Conexão de dados

Essa conexão será utilizada quando criarmos nosso próprio modelo dimensional.

1. Verifique se o Pentaho foi iniciado de forma adequada. Ele precisa estar inicializado na máquina para essa fase ocorra corretamente.
2. Antes de criar uma nova conexão copie o arquivo "biserver-ce/tomcat/common/lib/postgresql-8.2-504.jdbc3.jar" para a pasta "biserver-ce-3.5.0.stable/administration-console/jdbc/"
3. Inicie o serviço da Área Administrativa, como descrito acima.
4. Vá no menu Administration --> Database Connections e adicione uma nova conexão.

5. Preencha os campos
 1. Name: PostGreSQL
 2. URL:jdbc:postgresql://localhost:5432
 3. User Name: postgres
 4. Password: postgres

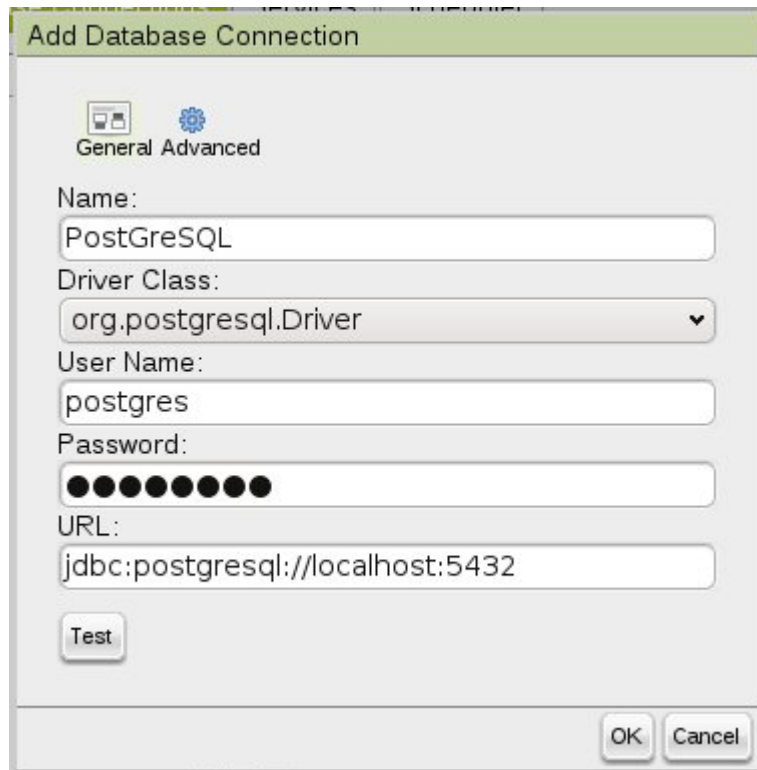


Figura 3 - Adicionar conexão com banco de dados

Definição da arquitetura

Agora vamos definir a arquitetura que será utilizada, de quais bases serão capturados os dados e em qual base esses dados serão persistidos. Também definiremos neste etapa o escopo do projeto e quais ferramentas serão utilizadas. A depender das ferramentas que você utilizar, será necessário seguir um caminho diferente do que vamos apresentar.

Construção de modelo Multidimensional

Existe uma série de ferramentas de modelagem em que é possível criar o modelo dimensional. Algumas serão listadas abaixo com os seus respectivos links:

Power-Architect (<http://www.sqlpower.ca/page/architect>)

DBDesigner (<http://fabforce.net/downloads.php>)

No trabalho foi utilizado o Power-Architect que é uma ferramenta Open Source (GPLv3). Além de ser possível construir o modelo multidimensional, ele gera o script SQL e o executa no banco.

Instalação e Configuração

1. Faça o download no link http://www.sqlpower.ca/page/architect_download. A versão que utilizamos é a 0.9.13.
2. Descompacte o arquivo em uma pasta de sua preferência.
3. Execute o architect.jar com o comando "java -jar architect.jar".
4. Clique no menu Connections --> Add Source Connection --> New connection...
5. Preencha os campos como descrito na imagem e clique em Test Connection. O usuário padrão é "postgres" e a senha "postgres".
6. Observe se a conexão foi bem sucedida.

7. Clique em OK.

Criação do modelo e geração da base de dados

1. Crie o modelo usando os botões no menu à direita. Este é o modelo que nós criamos e vamos utilizar neste tutorial.

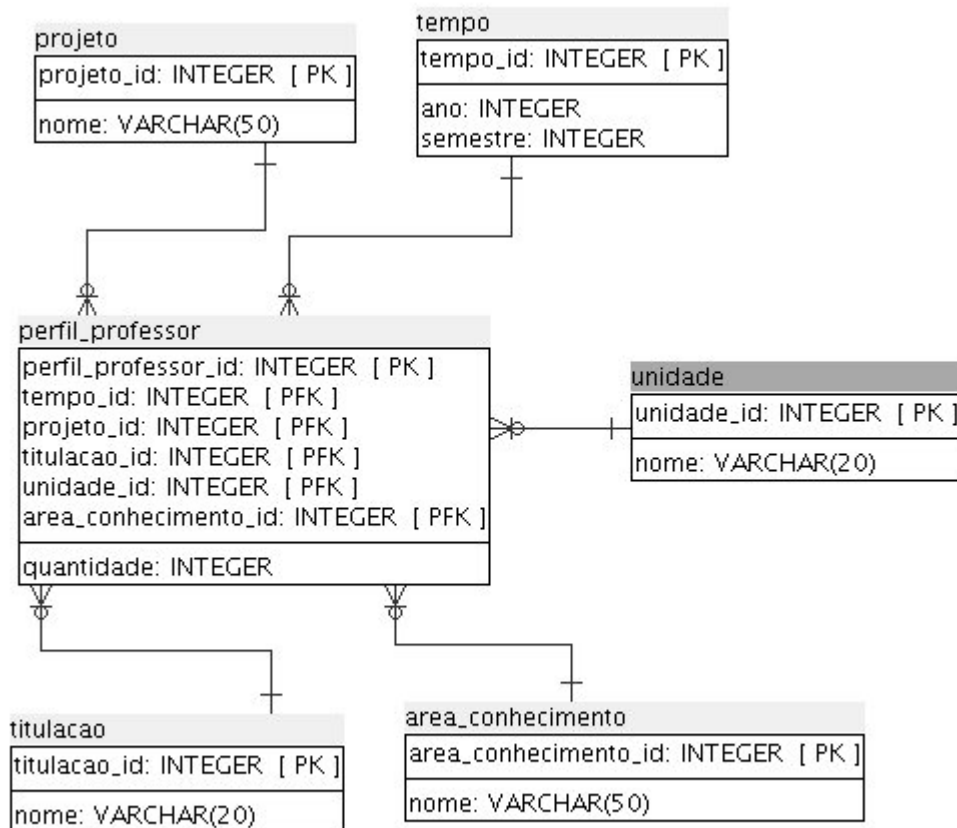


Figura 4 - Modelo dimensional

2. Clique no menu Tools --> Forward Engineer...

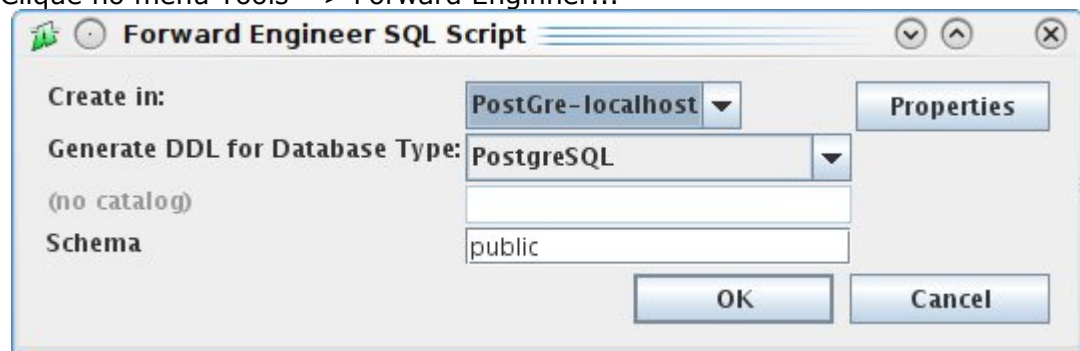


Figura 5 - Gerar script SQL a partir do modelo dimensional

3. Selecione a base criada. Preencha o Schema como "public" e clique em OK.
4. É possível copiar o SQL para ser executado em outra ferramenta ou executar através do próprio Power-Architect.

Criar o schema para identificação do cubo OLAP

O modelo do banco dimensional não é suficiente para o Pentaho identificar os cubos e dimensões. É necessário criar um arquivo XML descrevendo quais são os cubos e dimensões e suas respectivas tabelas associadas. Além disso, a partir desse XML, gerar um XACTION, que é o arquivo que irá ser identificado pelo Pentaho.

Entre as ferramentas que podem ser utilizadas nesta etapa estão:

- O próprio Power-architect;
- Schema Workbench (<http://sourceforge.net/projects/mondrian/files/>) e;
- Cube Dsigner.

Schema Workbench

Escolhemos o Schema Workbench, pois ele tem uma integração com o Pentaho e faz a publicação do xml automaticamente.

Funcionalidades

- Editor de esquema integrado.
- Testes das consultas **MDX** (Multi-dimensional Expressions - Linguagem criada pela Microsoft para definir consultas multi-dimensionais) nos esquemas e nas bases de dados.
- Visualização da estrutura de bases de dados.

Instalação e Configuração

1. A versão utilizada é a 3.2 stable. Efetue o download em (<http://sourceforge.net/projects/mondrian/files/>);
2. Descompacte o arquivo em uma pasta de sua preferência;
3. Para iniciar o Pentaho execute o arquivo chamado workbench.bat (Windows) ou workbench.sh (Linux).

Criando uma conexão com o Banco de dados

1. Vá em Tools --> Connection.
2. Preencher os campos
 1. Driver: org.postgresql.Driver.
 2. Connection URL: jdbc:postgresql://localhost:5432.
 3. User name e password: postgres.
3. Clique em Test Connection. Você deverá receber a mensagem: "Database Connection Sucessfull".
4. Clique em Accept.

Criando um Schema para identificação de um cubo

1. Vá em File --> New --> Shema
2. Clique no botão para inserir o cubo e adicione um nome que referenciará o cubo.



3. Clique com o botão esquerdo sobre o cubo e vá em Add Table. Preencha os campos corretamente com os dados da tabela fato definida no projeto do modelo

multidimensional.

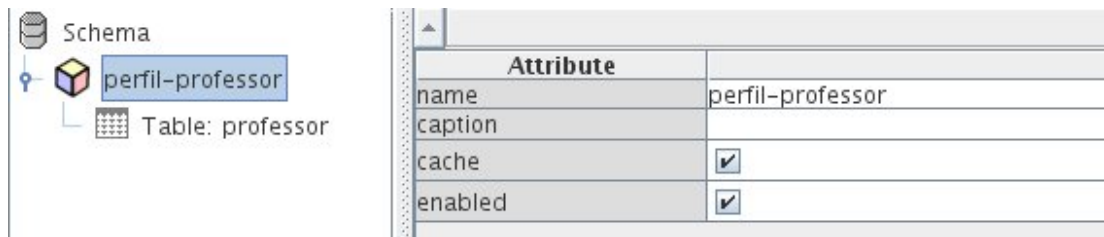


Figura 6 - Criação do cubo no Schema-Workbench

4. Clique com o botão direito do mouse sobre o cubo e vá em Add Measure. Essa entidade se referencia à métrica definida no modelo multidimensional. Preencha os campos
 1. Name: quantidade;
 2. Aggregator: sum;
 3. Collum: quantidade; (Nome da métrica que se encontra na tabela fato no banco de dados)
 4. dataType: Numeric
5. Adicione uma dimensão.



6. Indique qual a respectiva tabela para esta hierarquia clicando com o botão direito sobre a hierarquia e depois em Add Table. Adicione as informações da tabela de dimensão correspondente a que foi criada no banco de dados. Pode ser que o Workbench gere automaticamente uma seleção de campos para que seja possível selecionar um no name. No entanto, é necessário avaliar se o que está sendo gerado é realmente a nomenclatura certa da tabela (Inclusive, se atentando a detalhes de letras em maiúsculo e minúsculo).

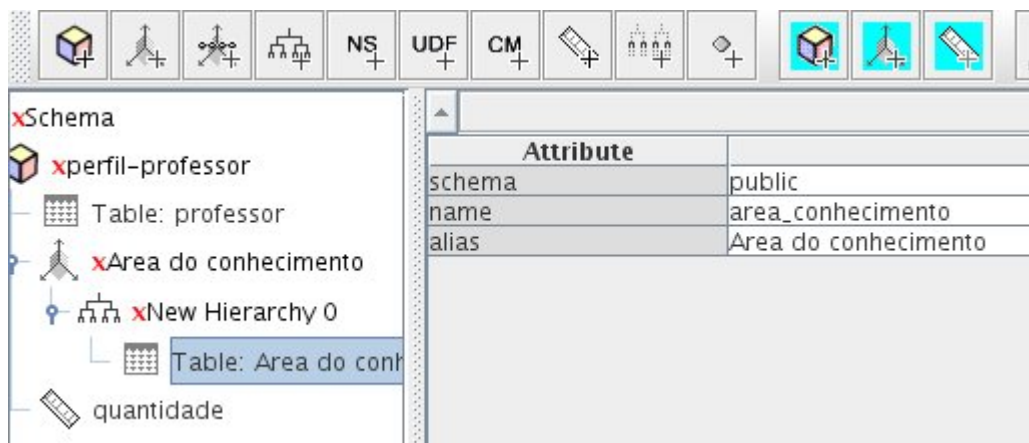


Figura 7 - Criando dimensões e hierarquias no cubo

7. Adicione um level para a Hierarquia criada. Esse level é equivalente à um atributo da tabela dimensão que deverá ser hierarquizado na dimensão de modo à serem gerados o drill down e drill up. Podem existir um ou mais levels em uma tabela dimensão.



8. Preencha os campos:

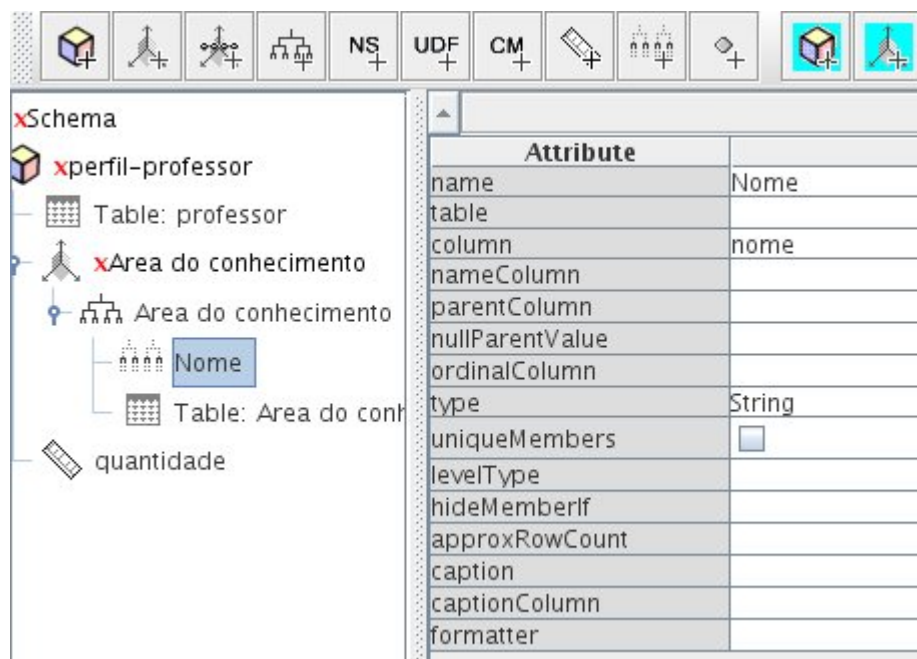


Figura 8 - Preenchendo informações da dimensão

9. **Repita os passo acima para as outras dimensões.**
10. No nosso modelo a dimensão tempo ficará assim

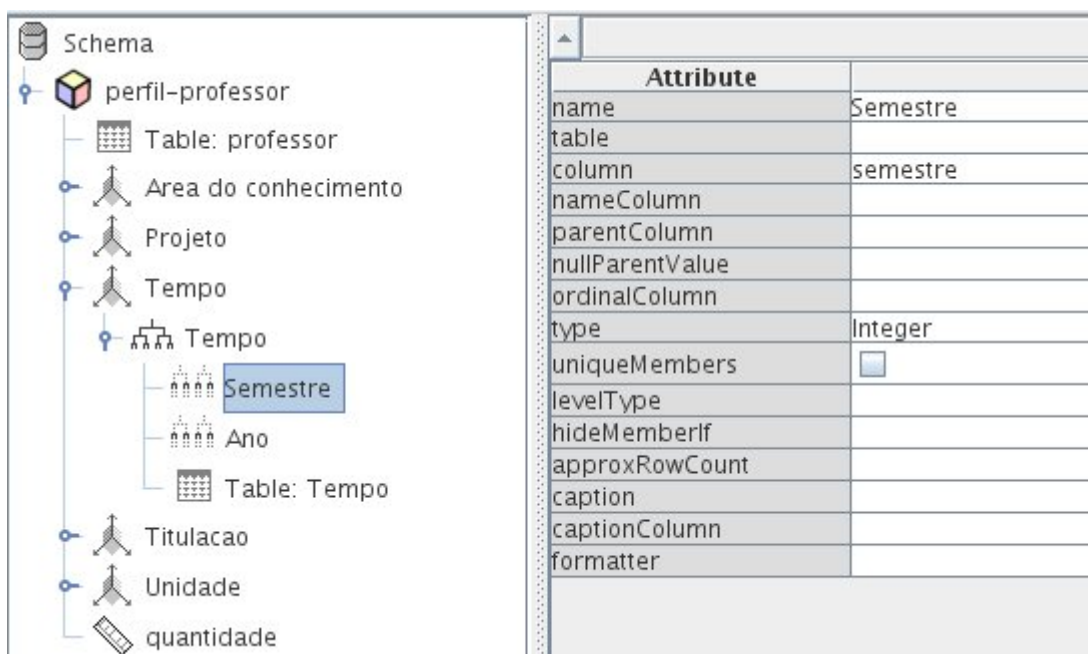


Figura 9 - Versão final do cubo

11. Observe a barra de status na parte de baixo do Schema-Workbench. Caso exista erros no modelo ali será exibido uma mensagem sinalizando.

Exportando o Schema criado para o Pentaho

1. Verifique se o Pentaho foi iniciado de forma adequada. Ele precisa estar inicializado na máquina para essa fase ocorra corretamente.
2. Abra o arquivo "biserver-ce/pentaho-solutions/system/publisher_config.xml"
3. É necessário adicionar uma senha:

```
<publisher-config>  
  <publisher-password>senha</publisher-password>  
</publisher-config>
```

4. Salve o Schema e clique em File --> Publish...
 1. Publish Password: "senha"
 2. user: "joe"
 3. password: "password"

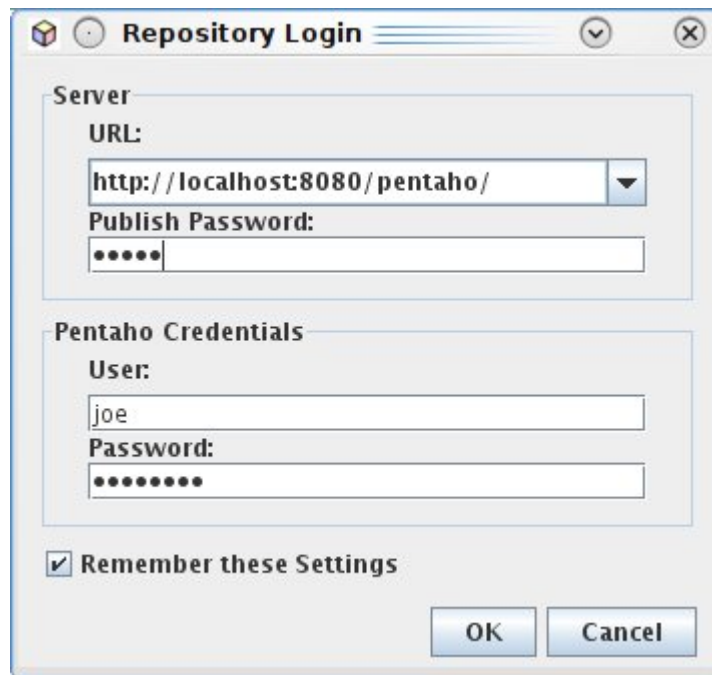


Figura 10 - Exportando o schema criado para o Pentaho

5. Clique em OK.

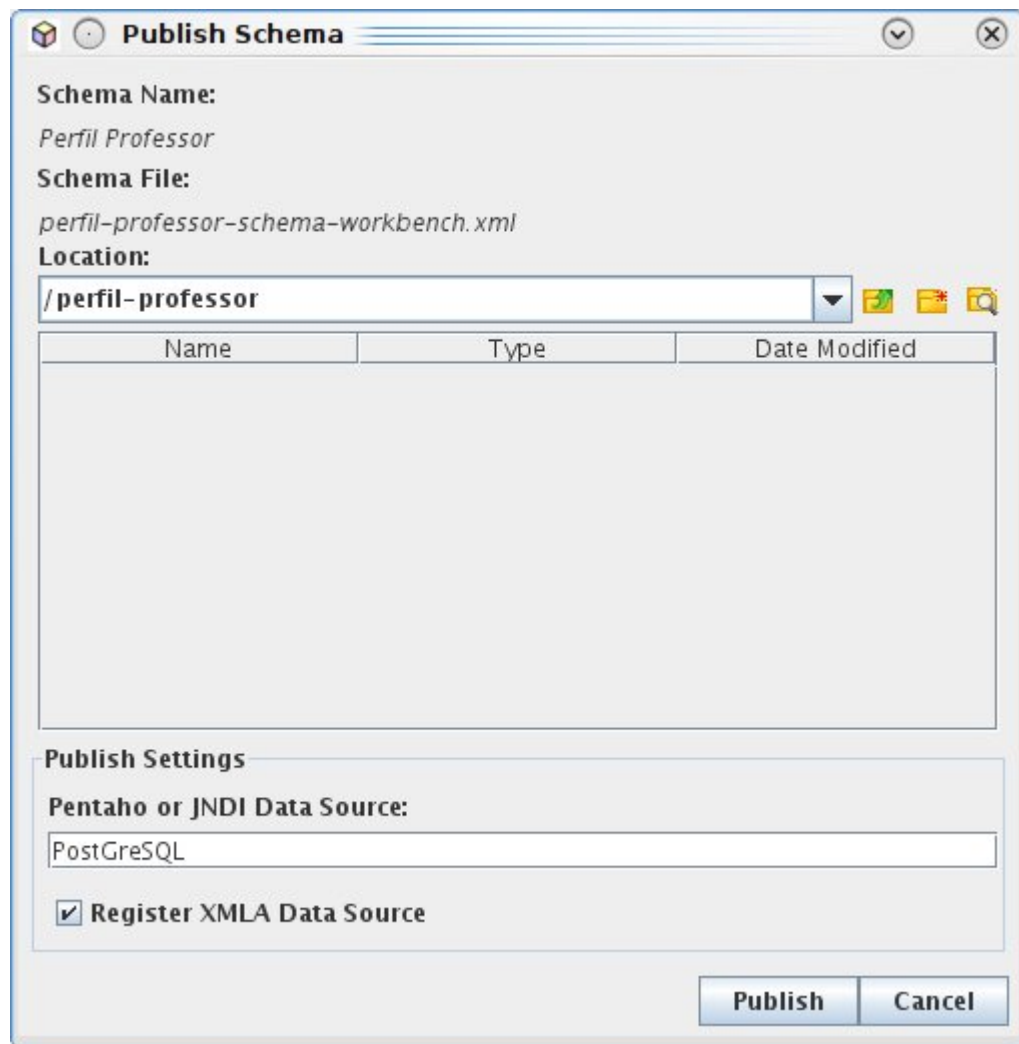


Figura 11 - Publicando schema no Pentaho

6. Crie uma nova pasta para o novo schema criado clicando no 2º botão ao lado do campo Location.
7. Defina em "Pentaho or JNDI Data Source" o nome da conexão de acesso à dados definida na fase [Criar nova Conexão de Dados](#).
8. Clique em Publish

Já é possível acessar o cubo criado. No Pentaho (navegador) clique em New Analyses View. Selecione o "Schema" (Ou o nome do Schema definido durante o processo de criação do schema no Schema Workbench) e o "Cube" criados. A janela deve ficar em branco, pois não há registros no banco de dados. Vamos inserir novos registros através do processo de ETL.

Processo de ETL (Extraction, Transformation and Load)

A ferramenta utilizada para fazer o processo de ETL é a Pentaho Data Integration - PDI, mais conhecida como Kettle. É uma ferramenta de código fonte aberto e sob a licença LGPL. Sendo uma ferramenta de ETL, ela foi construída para:

- Coletar dados de várias fontes distintas (Extract);

- Mover e modificar dados (Transform) enquanto limpa, desnormaliza, agrega e enriquece durante o processo;
- E carregar estes dados (Load) no destino final, que geralmente é um banco de dados enorme e que foi modelado dimensionalmente (data warehouse).

O PDI pode ser dividido em 3 aplicativos distintos:

- Spoon: É uma ferramenta gráfica usada para modelar as *transformations* e os *jobs*
 - *transformation*: modelo que captura o fluxo de dados de entrada e faz uma transformação até atingir o destino final.
 - *job*: modelo que tem entradas com o transformações (*transformations*), downloads FTP, etc. e saídas como envio de e-mails, por exemplo.
- Pan: Ferramenta de linha de comando usada para executar *transformations*.
- Kitchen: Ferramenta de linha de comando usada para executar *jobs*.

O PDI pode funcionar diretamente integrado a um repositório (um SGBD de sua preferência) ou utilizando arquivos. Neste documento iremos mostrar o seu uso conectado à um SGBD, assim não será necessário se preocupar com o armazenamento das transformações e jobs criados, isso será feito automaticamente pela ferramenta no banco de dados.

Download e Instalação

1. A versão utilizada é a 3.2 stable. Efetue o download em (<http://sourceforge.net/projects/pentaho/files/>);
2. Descompacte o arquivo em uma pasta de sua preferência;
3. Para iniciar o PDI execute o arquivo Kettle (Windows) ou spoon.sh (Linux);

Configuração

Agora vamos criar o repositório

1. Na tela inicial do PDI clique em Novo.
2. Na tela Repository Information clique em Novo para criar uma conexão com o banco;
3. Selecione o SGBD que será utilizado. Neste tutorial estamos utilizando o PostgreSQL.

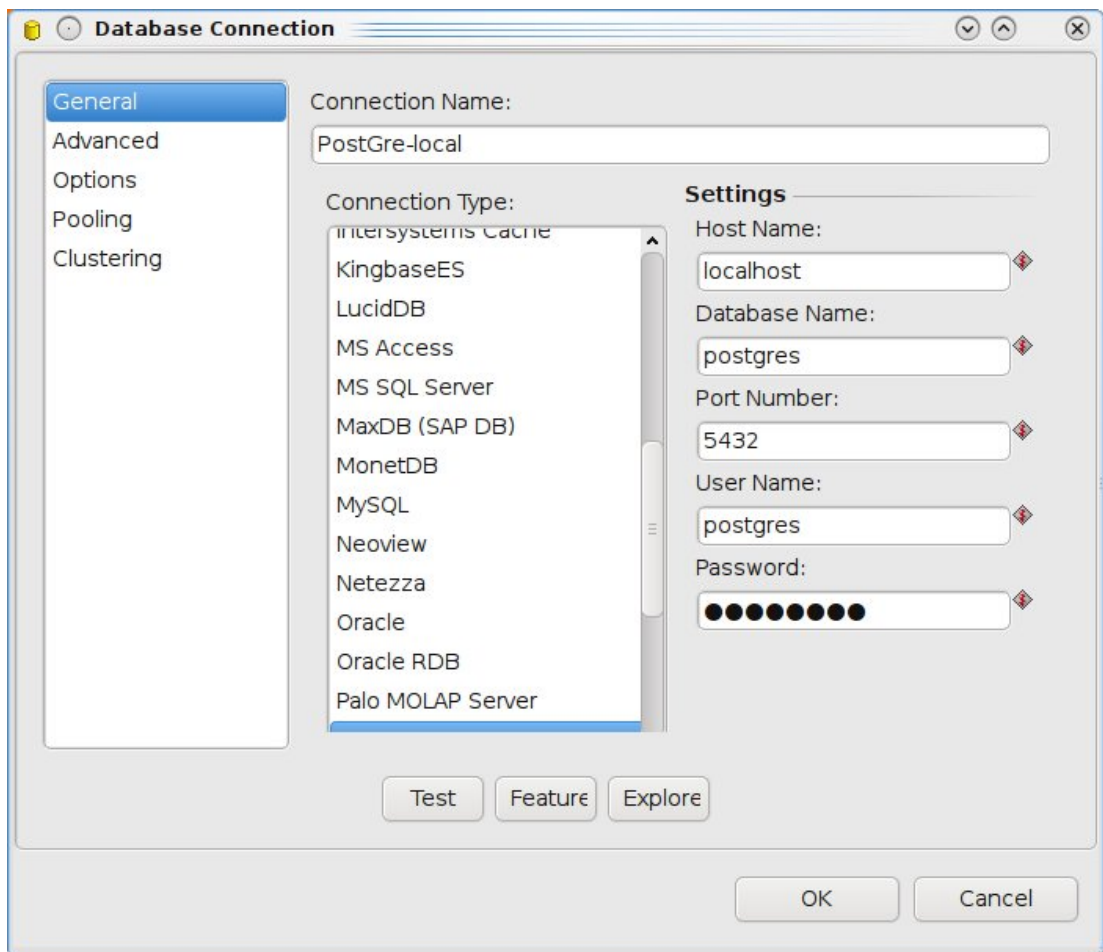


Figura 12 - Criando conexão com o banco de dados

1. Host Name: localhost.
2. Database Name: postgres.
3. Port Number: 5432 (padrão).
4. User name e password: postgres.
5. Clique em Test para ver se deu tudo certo.
4. Assegure que a conexão criada está selecionada, dê um nome ao novo repositório.

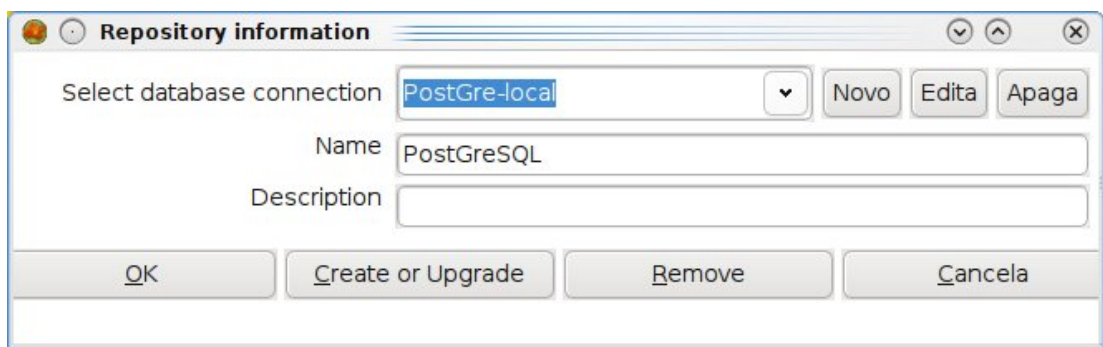


Figura 13 - Selecionando repositório

5. Clique em "Create or Update" e confirme. Neste momento o PDI irá criar as tabelas necessárias para o seu gerenciamento no banco de dados;
6. É perguntado então se você deseja ver/editar o script SQL que será executado, clique em Sim caso deseje, ou em Não para prosseguir;
7. Depois de criado o repositório clique em OK na tela Repository Information;

Selecione o repositório desejado, o usuário padrão é "admin" e a senha também é "admin", clique em OK. O PDI é então iniciado.

Utilização do PDI

Criação de uma transformação

1. Clique no menu em File --> New --> Transformation ou pressione Ctrl+N.
2. Clique no menu em Transformation --> Settings ou pressione Ctrl+T
3. Coloque um nome e uma descrição da transformação e clique em OK.

Conceitos importantes

Um Step é a unidade mínima em uma transformação. Existe uma grande variedade de Steps que são agrupados em categorias como Input e Output. Cada Step é definido para executar uma tarefa específica como a leitura de um parâmetro ou normalizar um conjunto de dados.

Um Hop é uma representação gráfica de um fluxo de dados entre dois Steps, com a origem e o destino. Um Hop só possui uma origem e um destino, mas mais de um Hop pode sair de um Step. Quando isso acontece, a saída de dados pode ser copiada ou distribuída em todo lugar. Da mesma forma, mais de um Hop pode chegar em um Step. Nesse caso, o Step precisa ter a capacidade de fundir as entradas de diferentes Steps, a fim de criar uma saída.

Construindo uma transformação. Utilizaremos um exemplo de transformação de informações de um arquivo de planilha para um banco de dados MySQL.

1. Clique em Design;
2. Input
 1. No menu lateral-esquerdo expanda o item Input;
 2. Arraste o "Excel Input" para tela de transformação;
 3. Clique duas vezes em cima do input para editá-lo;
 4. Coloque um nome no Step para fins de documentação e principalmente para localizá-lo mais facilmente depois;
 5. Selecione a planilha de interesse (botão Navega) e clique em Add. A planilha deve aparecer na lista Selected Files;
 6. Clique no menu em Sheets, clique em Get sheetname(s), selecione a planilha que deseja usar, clique no botão ">" e clique em OK;
 7. Insira a linha a partir do qual a entrada será lida (Start row);
 8. Clique no menu em Content, em Encoding selecione Windows-1250 caso seu documento esteja codificado no padrão windows;
 9. Se desejar marque a opção "Stop on empty row";
 10. Clique no menu em Fields e complete com as informações dos atributos que deseja mapear da planilha;
 11. Você pode clicar em "Preview Rows" para ver as linhas que serão retornadas;
 12. Clique em OK para sair do mode de edição;
3. Output
 1. Para adicionar uma saída, arraste o item Output --> Table output para a tela;
 2. Crie o Hop do input para o output: clique e arraste com o botão do meio do input ao output, ou com o botão esquerdo e pressionando Shift;
 3. Clique duas vezes em cima do output para editá-la.
 4. Dê um nome ao step;
 5. Crie uma nova conexão com o banco de dados destino (botão New);
 1. Selecione o banco de dados, coloque o nome do Host Name, database, porta, usuário e senha;

2. Clique em Test para testar a conexão e OK para finalizar;
3. Defina o schema destino e a tabela. Você pode clicar em Navega para selecioná-la;
6. Marque a opção "Specify Database Fields";
7. Na aba "Database Fields" clique em "Enter field Mapping" e relacione os campos provenientes do arquivo XLS com as colunas da tabela;
4. Executar a transformação: Transformation -> Run (ou aperte F9)
 1. Selecione "Detailed Logging" em "Log Level";
 2. Clique em Launch para executar a transformação.
5. Para inserir uma transformação entre o Input e o Output que possa ser executada no intermédio das duas
 1. Selecione Scripting
 2. Arraste Modified Java Script Value para a área de transformação
 3. Clique duas vezes em cima da instância que foi criada e defina um nome.
 4. Em Script 1 é possível escrever qualquer código JavaScript para conversão de dados da planilha para o banco.
 5. Clique em View no menu à esquerda.
 6. Clique com o botão direito em cima de Hops e clique em Novo.
 7. Selecione o Step de origem (O Input referente à planilha) e o de destino (Modified Java Script Value), selecione "Habilita Hop" e clique OK.
 8. Adicione um outro Hop.
 9. Selecione o Step de origem (Modified Java Script Value) e o de destino (Output referente ao banco), selecione "Habilita Hop" e clique OK.
 10. Clique duas vezes em cima do Output.
 11. Clique em Database fields e clique em Get fields.
 12. Clique em Enter field mapping, selecione os atributos de origem e destino, clique em Add e em OK.
 13. Clique em Ok novamente

Análise do Cubo

O Analysis View utiliza como cliente o JPivot e a engine do Mondrian para servidor OLAP. Interface final do usuário, o software provê os recursos necessários para o gerenciamento de dados em um DW. O Mondrian é configurado através de esquemas, que são arquivos XML.

Abrindo Schema criado anteriormente

1. Na tela inicial do Pentaho clique em New Analysis View.
2. Selecione o Schema que criou
3. Irá abrir a tela de entrada do Analysis com o cubo modelado

Manual e utilização do Mondrian, feito por alunos da UFMG:

http://www.libertasbr.org.br/nagios-dw/docs/manual_mondrian.pdf

O JPivot é um visualizar um pouco difícil de ser manipulado por usuários finais ainda e, por isso, já está surgindo uma nova iniciativa para substituí-lo. A nova ferramenta é denominada PAT (Pentaho Analysis Tool). É possível obter maiores informações em:

<http://wiki.pentaho.com/display/COM/Pentaho+Analysis+Tool>

Adicionando Usuários

Para prover recursos de segurança, de maneira que o acesso aos dados estratégicos da organização possa ser restringido, o Pentaho fornece um gerenciador de usuários e papéis que pode ser configurado na [Área Administrativa](#).

É possível gerenciar os usuários e suas permissões de acesso, criar repositórios para o desenvolvimento de soluções de BI com ajuda do Pentaho, gerenciar os serviços do servidor do Pentaho e gerenciar a execução automática de tarefas.

Para criar um usuário:

1. No console de administração, acesse **Administration > Users & Roles**.
2. Caso não esteja na seção de usuários, clique no ícone **Users**.
3. Clique no sinal de mais(+), próximo a **Users**.
4. No painel de detalhes, digite o nome do usuário, senha, confirmação da senha e descrição (**User Name, Password, Password Confirmation, and Description**).
5. Uma vez finalizado a inserção dos dados acima, clique em OK e o novo usuário estará acessível via a lista de usuários.

Editando Usuários

Uma vez criado um usuário, é possível efetuar a atualização de suas informações através dos passos abaixo:

1. No console de administração, acesse **Administration > Users & Roles**.
2. Selecione o usuário cuja informação se deseja editar.
3. No painel de detalhes, edite os dados do usuário conforme desejado.
4. Após a edição, clique no botão atualizar (**Update**).

Deletando Usuários

É possível a exclusão de acesso de um usuário, através dos passos a seguir:

1. No painel de administração, acesse **Administration > Users & Roles**.
2. Na lista de usuários, selecione um usuário.
3. Clique no sinal de menos(-), próximo a Users, uma mensagem de confirmação irá aparecer.
4. Clique em OK para confirmar a exclusão do usuário.

Adicionando e excluindo Papéis

A adição de papéis consiste na criação de políticas de restrição de acesso aos usuários. Os passos para criação destes papéis são:

1. No painel de administração, acesse **Administration > Users & Roles**.
2. Clique no ícone papéis(Roles).
3. Clique no sinal de mais(+) ou menos(-), para adicionar ou excluir um papel a lista de papéis disponíveis, respectivamente.
4. Caso deseje adicionar, é necessário digitar o nome do papel e descrição. Além disso, é possível, na inserção, atribuir aos usuários este novo papel.
5. Clique em Ok para confirmar a opção desejada.

Personalização do Pentaho

Para editar a tela default de login do pentaho edite o arquivo PUCLogin.jsp

Local do arquivo : biserver-ce\tomcat\webapps\pentaho\jsp

Para editar a página central edite o arquivo launch.jsp

Local : biserver-ce\tomcat\webapps\pentaho\mantle\launch

Para editar a página de carregamento edite o arquivo DebugMantle.html

Local : biserver-ce\tomcat\webapps\pentaho\mantle

Para editar a página de carregamento do console de administrador edite o arquivo:
PentahoAdminConsole.html

Local : administration-console\www\org.pentaho.pac.PentahoAdminConsole

Tradução do Pentaho

- Faça o download do pacote de tradução em:

http://brazilianportuguesetranslationofpentaho.googlecode.com/files/prof.coruja-pentaho-traducao3.0-stable_pt_BR.zip

Existem arquivos *.properties*, e para cada localidade o sistema automaticamente escolhe a versão correta. Por exemplo, se o navegador está configurado para o idioma Inglês, o sistema utiliza o conteúdo do arquivo *messages_en.properties*. Se for Português, será utilizado o conteúdo do arquivo *messages_pt_BR.properties*. Quando um arquivo específico para o idioma não for encontrado, o arquivo *messages.properties* será utilizado.

É possível traduzir inclusive seus cubos, relatórios, etc. Basta criar um arquivo igual contendo a tradução e a nomenclatura correta.

Para traduzir, substitua os arquivos da tradução que você baixou e coloque em seu Pentaho.

Arquivos da tradução do Sistema encontram-se em dois lugares:

biserver-ce\tomcat\webapps\pentaho\mantleLogin e

biserver-ce\tomcat\webapps\pentaho\mantle

Outros arquivos *.properties* encontram-se em todo o sistema principalmente na pasta pentaho-solutions e o processo de tradução é igual.

Como já foi mencionado é possível utilizar esse padrão em todo o sistema, todo o sistema pode ter arquivos *.properties* com o idioma necessário.

Estes arquivos precisam ser salvos no formato UTF-8

Outras ferramentas

Além destas, existem outras ferramentas para análise de BI.

Pentaho Analysis Tool (PAT)

O PAT é uma ferramenta de análises que está sendo criada pela comunidade de usuários do Pentaho para substituir o JPivot. Esta ferramenta se encontra em bom estágio de desenvolvimento e já na versão 0.5 (atual) consegue gerar análises via *drag and drop* (arrastar e largar) e via código MDX, além de outras pequenas funcionalidades.

Table

Dimensions

- Unused**
- Positions
- Rows**
- Region
- Department
- Columns**
- Measures
- Filter**
- Properties**
 - Show MDX
 - Show Parent
 - Show Filters
 - Show Properties
 - Hide Blank Cells
 - Pivot/Swap Axis
 - Layout

Region	Department	Actual
All Regions	Executive Management	143,639,982.00
	Finance	6,299,022.00
	Human Resource	12,224,220.00
	Marketing & Communication	13,075,463.00
	Product Development	13,910,753.00
Central	Product Development	10,644,102.00
	Professional Services	76,317,649.00
	Sales	11,168,773.00
	Executive Management	37,893,162.00
	Finance	1,776,282.00
Eastern	Finance	3,106,680.00
	Human Resource	3,438,863.00
	Marketing & Communication	3,590,423.00
	Product Development	2,997,702.00
	Professional Services	20,068,039.00
Southern	Sales	2,915,173.00
	Executive Management	35,248,940.00
	Finance	1,507,580.00
	Human Resource	3,039,180.00
	Marketing & Communication	3,212,200.00
All Regions	Marketing & Communication	3,440,110.00
	Product Development	2,548,800.00
	Professional Services	18,749,870.00
	Sales	2,751,200.00
	Executive Management	35,248,940.00
Central	Executive Management	1,507,580.00
	Finance	3,039,180.00
	Human Resource	3,212,200.00
	Marketing & Communication	3,440,110.00
	Product Development	2,548,800.00
Eastern	Professional Services	18,749,870.00
	Sales	2,751,200.00
	Executive Management	1,507,580.00
	Finance	3,039,180.00
	Human Resource	3,212,200.00
Southern	Marketing & Communication	3,440,110.00
	Product Development	2,548,800.00
	Professional Services	18,749,870.00
	Sales	2,751,200.00
	Executive Management	1,507,580.00

Chart

Figura 14 - Imagem do PAT (PAT, 2009)

Pentaho Analyser

Recentemente foi adicionada à versão Enterprise do Pentaho a ferramenta Pentaho Analyser que é basicamente uma versão completa funcional do PAT. Ela permite operações *drag and drop* (arrastar e largar), cálculos definidos pelo usuário, filtros com um clique, exportação para o Excel, etc.

File View Help

KPI Dashboard **Untitled**

Hide Field List **Show Filters** **Hide Field Layout** **Reset** **More**

Unsaved Report **View As:** **Table**

No Filter in use **Your report is ready.** **Rows: 28** **Columns: 2**

Row Labels **Department** **Region**

Col Headers **Actual** **Budget**

Department	Region	Actual	Budget
Executive Management	Central	1,776,282.00	2,043,642.00
	Eastern	1,507,580.00	1,483,508.00
	Southern	1,507,580.00	1,523,508.00
	Western	1,507,580.00	1,443,508.00
Finance	Central	3,106,680.00	3,067,361.00
	Eastern	3,039,180.00	3,010,015.00
	Southern	3,039,180.00	3,010,015.00
	Western	3,039,180.00	3,000,015.00
Human Resource	Central	3,438,863.00	3,414,295.00
	Eastern	3,212,200.00	3,195,682.00
	Southern	3,212,200.00	3,195,682.00
	Western	3,212,200.00	3,183,682.00
Marketing & Communication	Central	3,590,423.00	3,582,552.00
	Eastern	3,440,110.00	3,383,905.00
	Southern	3,440,110.00	3,401,905.00
	Western	3,440,110.00	3,401,905.00
Product Development	Central	2,997,702.00	3,159,180.00
	Eastern	2,548,800.00	2,530,477.00
	Southern	2,548,800.00	2,548,477.00
	Western	2,548,800.00	2,548,477.00
Professional Services	Central	20,068,039.00	20,400,000.00
	Eastern	18,749,870.00	19,100,000.00
	Southern	18,749,870.00	18,390,000.00
	Western	18,749,870.00	18,208,206.00

Available fields (6) for: Quadrant Analysis

Find: **View**

- Department**
 - Department
- Measures**
 - Actual
 - Budget
 - Variance
- Positions**
 - Positions
- Region**
 - Region

Figura 15 - Exemplo retirado da versão Demo do site do Pentaho (PENTAHO, 2009)

Pentaho Report Designer (PRD)

Além das análises o Pentaho também pode gerar relatórios para visualização mais detalhada dos dados. A ferramenta padrão para criação de relatórios é o Pentaho Report Designer (PRD).

O PRD é uma ferramenta com interface intuitiva e de fácil uso, com apenas alguns cliques já é possível criar um relatório simples. A interface de ajuste de *layout* de relatórios lembra muito a interface de ajuste de relatórios do Microsoft Access e pode ser vista abaixo:

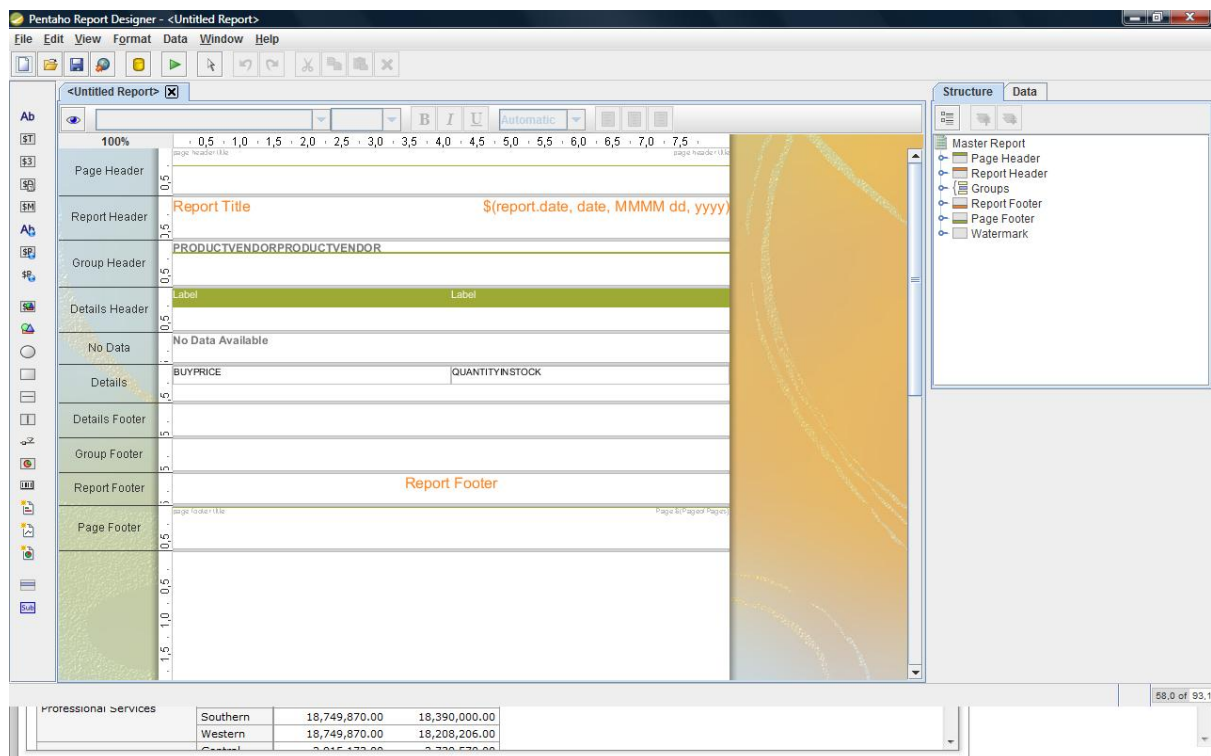


Figura 16 - Pentaho Report Designer (PENTAHO, 2009)

Para efetuar a instalação do PRD será necessário:

- Efetuar o download da ferramenta no repositório: <http://sourceforge.net/projects/pentaho/files/>. Nesse trabalho foi utilizada a versão 3.5 stable
- Extraia o conteúdo do arquivo compactado que foi baixado
- Inicie a ferramenta executando report-designer.bat (Windows) ou report-designer.sh (Linux).
- Para um primeiro exemplo na ferramenta, vamos utilizar o mecanismo de wizard para facilitar o processo de criação do relatório. Clique em File --> Report Wizard ou pressione Ctrl+Shift+N.
- É possível identificar à direita da janela que irá abrir a etapa que você está. A primeira é o Look and Feel. Escolha nesse momento como vai ser o visual do seu relatório. Existem alguns tipos prontos e, é possível, através da seleção do tipo report Document, de um modelo produzido pelo usuário. Após a escolha clique Next.

- Na etapa Data Source and Query vai ser identificado o contexto do relatório. É possível fazer essa identificação através de vários tipos que podem ser visualizados clicando no "+" localizado no canto direito superior. Nesse caso, usaremos o modelo que gerados para visualizar os dados no Analysis. Selecione Pentaho Analysis.
- Clique em Browse e selecione o arquivo .xml que tem o modelo do schema que foi publicado no Analysis. Normalmente, ele está localizado na pasta do seu projeto em \biserver-ce\pentaho-solutions.
- Crie uma conexão clicando no "+" ao lado de connections.
- Coloque um nome na conexão, indique o tipo da conexão, selecione em Access "Native (JDBC)" e coloque os dados referentes à sua conexão do lado direito.
- Teste a conexão clicando em "Test" e se tudo estiver ok, clique em OK.
- Crie uma query e adicione uma query usando a linguagem MDX correspondente à consulta que deseja utilizar no seu relatório. É possível gerar essa consulta através do Pentaho Analysis. Clique OK.
- Na etapa Layout Step será possível indicar quais campos da consulta que foi selecionada na etapa anterior que serão utilizados no relatório. Indique quais campos serão os agrupadores e os itens.
- Na etapa Format Step é possível configurar melhores nomes para os campos e itens do relatório.
- Clique em Finish e será gerado um relatório baseado nas informações respondidas no wizard.
- É possível manualmente efetivar outras monificações.

Web Ad Hoc Query and Reporting Service (WAQR)

O Pentaho também oferece em seu console de usuário a opção de geração de relatórios Ad Hoc através do serviço WAQR, sem que seja necessária a utilização do PRD. Relatórios gerados pelo WAQR, no entanto, devem seguir templates pré-estabelecidos, criados utilizando o PRD como visto na figura abaixo:

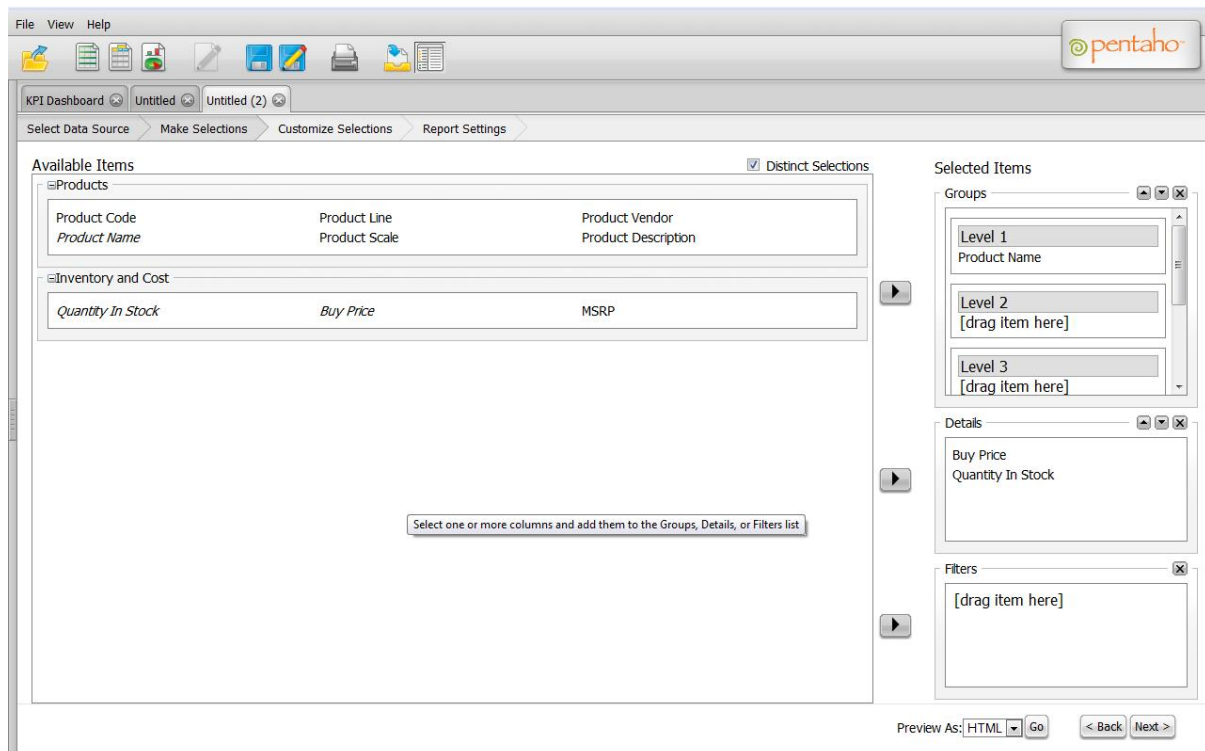


Figura 17 - Exemplo de um relatório sendo gerado utilizando o WAQR (PENTAHO, 2009)

Community Dashboard Framework

Dashboards são coleções de gráficos e/ou relatórios geralmente agrupados com um mesmo tema para facilitar a visualização e compreensão das informações pelos usuários. O Pentaho Community Edition possui um framework de criação de dashboards chamado Community Dashboard Framework (CDF) que facilita bastante a criação de dashboards. O CDF, porém requer de seus desenvolvedores conhecimento de HTML, JavaScript, JQuery e CSS, o que pode dificultar a criação de dashboards por desenvolvedores menos experientes ou até mesmo por usuários comuns.

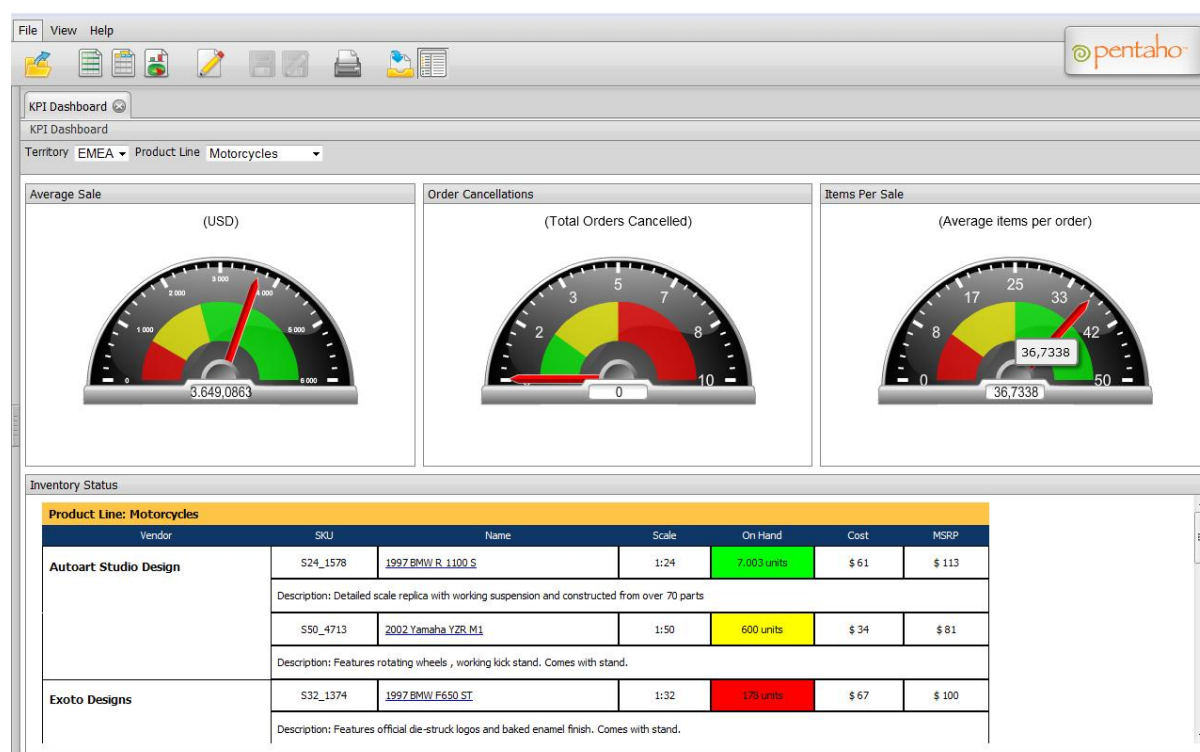


Figura 18 - Exemplo Dashboard (PENTAHO, 2009)

Com o intuito de resolver este problema, foi desenvolvido pela comunidade o CDF Dashboard Editor (CDF-DE). O CDF-DE tem uma interface simples e intuitiva com operações *drag and drop* (arrastar e soltar) possibilitando a criação de dashboards com muito mais facilidade.

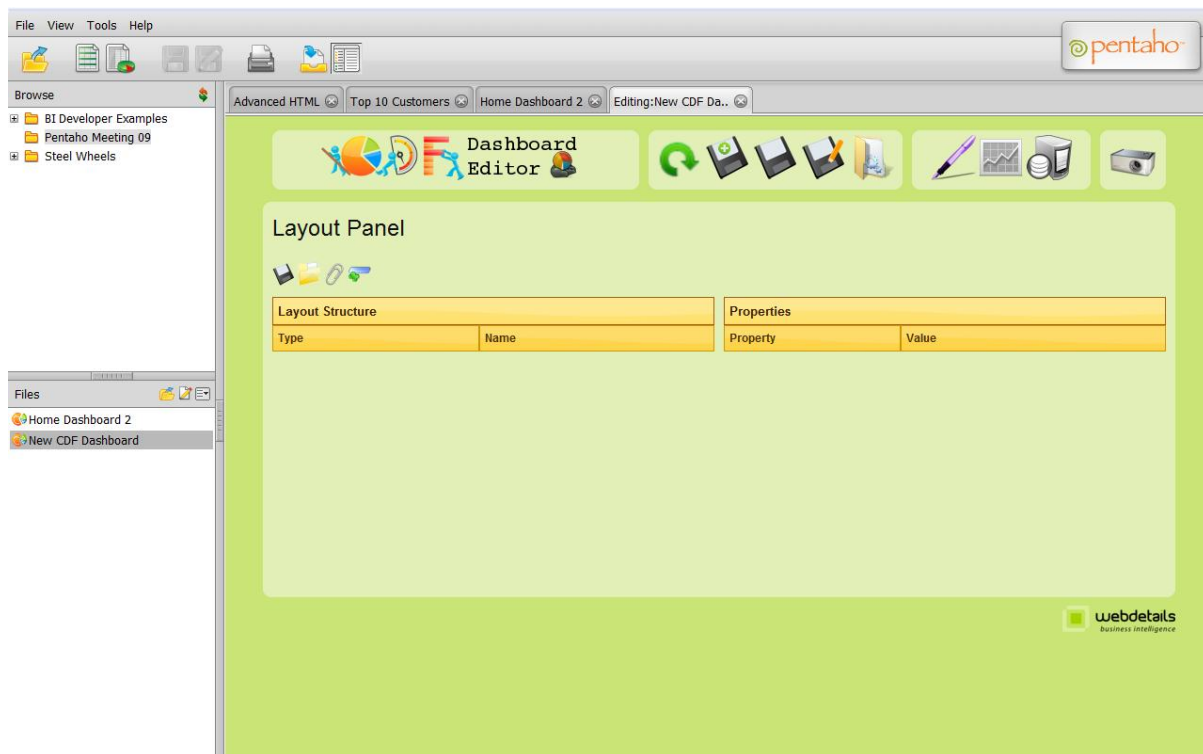


Figura 19 - Exemplo da interface de criação de dashboards do CDF-DE

No nosso trabalho, utilizamos o CDF-DE para efetuar a geração dos Dashboards. É possível efetuar a instalação seguindo os seguintes passos:

- Baixe o arquivo: `cdf-de_0.2.tar.bz2`. O arquivo encontra-se disponível para download em: <http://code.google.com/p/cdf-de/downloads/list>
- Descompacte o arquivo `cdf-de_0.2.tar.bz2`.
- Copie a pasta `system` para: `biserver-ce\pentaho-solutions`.
- Copie a pasta `pentahomeeting09` para: `biserver-ce\pentaho-solutions`.
- Copie a pasta `cdf` para: `biserver-ce\tomcat\webapps\pentaho`.
- Reinicie o servidor do Pentaho

Para efetuar a geração de dashboards é necessário seguir os seguintes passos:

- Inicie o Pentaho e vá em `File --> New --> New Dashboard`.
- Clique em `Datasources`.
- Vá em `Wizards -> OLAP Chart wizard`.
- Selecione o `schema` e o `cubo`.
- Serão geradas automaticamente as dimensões. Dessa forma vai ser possível clicar e arrastar para definir as linhas, colunas e o que vai ser filtrado no dashboard.
- Após definir as informações que serão mostradas nos dashboards, selecione na coluna à direita as configurações visuais do dashboard.
- Verifique se na área de visualização do Dashboard a opção *"Preview Area"* está selecionada.

Créditos

Direitos Autorais Reservados (c)
Universidade Federal da Bahia
Instituto de Matemática

Departamento de Ciência da Computação

Manuele Ferreira (manueleferreira@gmail.com)
Gustavo Ramos (gustavoramos00@gmail.com)
Luis Bernardo (luber.braga@gmail.com)
Robson Silva (robson.rss@gmail.com)
Bruno Maior (fox.ssa@gmail.com)

Referências

CONSTRUCAO DE DATA WAREHOUSE PARA PEQUENAS E MÉDIAS EMPRESAS USANDO SOFTWARE LIVRE, MARCELO FEIJÓ VARGAS, 2008, LAGES (SC)
[Publishing an Analysis Schema Using Schema Workbench](#)
<http://wiki.pentaho.com/display/EAI/Getting+Started>
[Latest Pentaho Data Integration \(aka Kettle\) Documentation](#)
<http://wiki.pentaho.com/display/COM/Pentaho+Data+Integration+Articles>
<http://rpbouman.blogspot.com/2006/06/pentaho-data-integration-kettle-turns.html>
<http://wiki.pentaho.com/display/EAI/Special+database+issues+and+experiences>
<http://blog.professorcoruja.com/2009/03/entendendo-o-pentaho-faq.html>
<http://blog.professorcoruja.com/2009/06/traducao-pentaho-bi-server-20-e-30-para.html>
[Manual da ferramenta Mondrian Schema Workbench](#)
[MANUAL DE INSTALAÇÃO DO PENTAH0 E PERSONALIZAÇÃO](#)
<http://blog.professorcoruja.com/>
<http://www.fusioncharts.com/free/>
<http://wiki.pentaho.com/display/Reporting/1.+Quick+Start>
<http://blog.professorcoruja.com/2009/10/como-instalar-o-cdf-dashboard-editor.html>

Esta documentação é livre; é possível redistribuí-la e/ou modi cá-la sob os termos da Licença Pública Geral GNU conforme publicada pela Free Software Foundation; tanto na sua versão 2, como qualquer versão posterior (a seu critério) .

A distribuição desta documentação é realizada com o objetivo de que ela auxilie os iniciantes da ferramenta e que não tenha qualquer garantia implícita de comerciabilidade ou adequação a uma finalidade específica.



Consulte a Licença Pública Geral do GNU para mais detalhes.
<http://creativecommons.org/licenses/GPL/2.0/>
<http://creativecommons.org/licenses/GPL/2.0/deed.pt>