

PENTAHO SEBAGAI SOLUSI MASALAH PENGOLAHAN DATABASE

(Pentaho as a Solution of Database Processing Problems)

Nurtriana Hidayati

Fakultas Teknologi Informasi dan Komunikasi Universitas Semarang

Abstract

Information is the result of data processing plays an important role in an organization, especially in decision-making process. Pentaho application of Intelligent Business Products is one of the technologies for collecting, storing, analyzing, and providing access to data to help enterprise users make better business decisions. Pentaho has a function as reporting, analysis, dashboards, data integration (ETL) and data mining. Pentaho is better to manage large and complex data and be able to complete the functional organization.

Keywords : *Pentaho, Database*

1. PENDAHULUAN

Informasi adalah fondasi dari tiap keputusan bisnis penting yang dilakukan. Dalam tiap perusahaan informasi merupakan hal yang perlu diatur dengan konsep yang benar sesuai proses bisnis yang terjadi. Pembuatan dan penerapan sistem ERP mengintegrasikan semua area fungsional dalam perusahaan. Sistem ERP merupakan tulang punggung dari kebanyakan perusahaan yang mengintegrasikan aplikasi back-office seperti aplikasi keuangan, pembelian, HR, inventory dan lainnya.

Namun, walau suatu sistem ERP dapat mengintegrasikan semua data transaksi bisnis kedalam masing-masing tabel master untuk perencanaan perusahaan. ERP bukanlah sistem untuk melakukan analisa data dan proses pendukung keputusan yang pengimplementasiannya biasa dikenal sebagai decision support system.

Fungsionalitas pendukung keputusan sangat penting untuk suatu perusahaan karena membantu perusahaan merencanakan strategi kedepan dan mengoptimalkan waktu dalam pembuatan keputusan. Suatu sistem Bisnis Intelijen (BI) dapat menyediakan pemetaan data dari suatu sistem ERP dan menganalisa skenario bisnis yang berjalan karena suatu BI melakukan pengkoleksi-

an data yang didapat dari sistem ERP, CRM atau sistem lainnya dan menghasilkan berbagai macam bentuk analisa. BI menyediakan analisa yang lebih bermakna. Walaupun laporan-laporan operasional yang disediakan ERP memberikan status kondisi bisnis saat ini, hal tersebut belum memenuhi keinginan para manager akan adanya ad hoc, forecasting dan exceptional report lainnya. BI juga dipakai untuk mengoptimasi investasi dari ERP. Suatu sistem ERP yang diintegrasikan dengan aplikasi bisnis intelijen secara kontinu dapat meningkatkan kelebihan kompetitifnya. Sebagai contoh, hasil dari solusi BI dapat membuat pegawai pembelian menemukan pattern harga, dimana dapat menyebabkan perusahaan mendapatkan harga yang lebih baik dengan mengubah proses pembelian. Penemuan-penemuan hasil analisa dari info yang diberikan oleh BI inilah yang dapat meningkatkan sistem dalam ERP.

2. KAJIAN PUSTAKA

Kajian pustaka ini akan menampilkan beberapa kajian teori yang dipakai dalam membangun aplikasi *Bisnis Inteligent* (BI), baik dalam pembangunan BI secara internal, daftar referensi dari semua jenis referensi seperti buku, *jurnal papers*, artikel, disertasi, tesis, skripsi, *hand outs*, *laboratory*

manuals, dan karya ilmiah lainnya serta modul-modul yang terkait dalam proses pengintegrasian aplikasi BI dengan ERP.

2.1 Bisnis Inteligent

Bisnis Intelijen (BI), merupakan kategori yang luas dari aplikasi dan teknologi untuk mengumpulkan, menyimpan, menganalisis, dan menyediakan akses ke data untuk membantu pengguna enterprise membuat keputusan bisnis yang lebih baik. Secara garis besar teknik yang dipakai dalam BI adalah teknik untuk mengidentifikasi, mengekstraksi, dan menganalisa data-data dari bisnis tertentu berdasarkan kategori-kategori yang mempengaruhi, contoh: menganalisa pendapatan penjualan (sales revenue) berdasarkan produk yang dijual dan cabang.

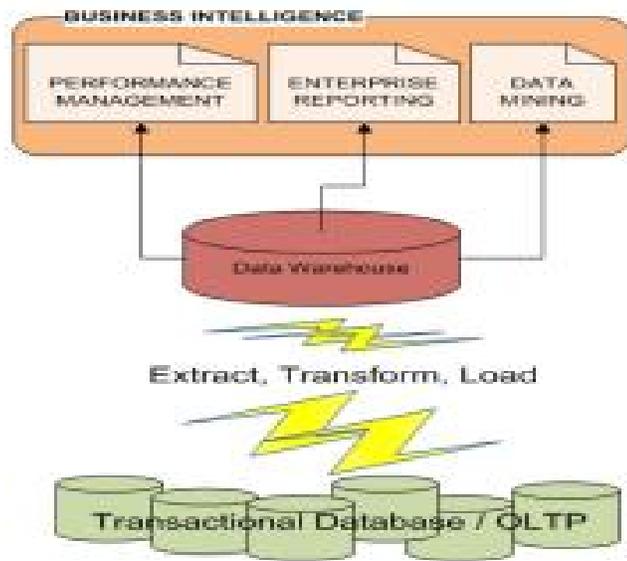
Teknologi BI menyediakan akses kepada data-data lampau, saat ini, dan data prediksi dari operasi-operasi bisnis. Adapun fungsionalitas umum dari BI adalah

1. Query dan Analysis yang dilakukan dengan cara :
 - a. Ad hoc reporting
 - b. analysis tools
2. Reporting yang dilakukan dengan cara :
 - a. Dashboard
 - b. Trend Line Analysis

3. Plan, Budget and Forecast yang dilakukan dengan cara :
 - a. Plan vs Actual
 - b. Performance Management yang dilakukan dengan cara:
 - a. KPI
 - b. Scorecards
4. Berdasar Waktu yang dilakukan dengan cara :
 - a. Pengolahan data untuk menjawab pertanyaan-pertanyaan bisnis

Data yang dipakai dalam BI sering merupakan Data Warehouse atau Data Mart. Namun, suatu BI tidak selalu membutuhkan suatu data warehouse, dan suatu data warehouse tidak selalu digunakan untuk keperluan BI.

Selanjutnya data di data warehouse diproses menggunakan berbagai analisis statistik dalam proses data mining, serta analisa OLAP, sehingga didapat berbagai kecenderungan atau pattern dari data yang merupakan report dari BI. Yang disajikan dengan tool BI yang telah tersedia. Report tersebut mampu membantu organisasi dalam pengambilan keputusan bisnis yang diambil berdasarkan fakta-fakta aktual. Berikut adalah gambaran mengenai konsep BI :



Gambar 1 Konsep Business Intelligence

2.2 Pentaho

Pentaho merupakan open source BI sekaligus menyediakan platform dan program untuk membangun suatu aplikasi BI. Pentaho terdiri dari banyak program komputer yang bekerja secara bersamaan dan menyediakan solusi bisnis intelijen.

Karena terdiri dari berbagai komponen, dimana ada komponen level bawah, namun juga ada komponen yang menyediakan fungsionalitas level atas yang biasanya mengandalkan fungsionalitas yang ditawarkan oleh komponen level bawah, keseluruhan koleksi komponen dalam Pentaho dapat dilihat seperti stack, yang mana semakin tinggi levelnya maka akan semakin dekat dengan end-user.

Keseluruhan produk pembangun aplikasi BI biasa disebut dengan Pentaho BI Suite yang terdiri dari beberapa produk, dimana masing-masingnya memiliki spesifikasi fungsi tersendiri, antara lain :

1. Pentaho Reporting :
 - a. Pentaho Report Designer
 - b. Pentaho Metadata Editor
2. Pentaho Analysis :
 - a. Pentaho Schema Workbench
3. Pentaho Dashboards :
 - a. Pentaho Design Studio
4. Pentaho Data Integration (ETL) :
 - a. Pentaho Data Integration
5. Pentaho Data Mining :
 - a. Weka

3. STUDI PUSTAKA

3.1 Pentaho Reporting

Pentaho Reporting adalah suatu koleksi proyek open source untuk merancang, menghasilkan dan mendistribusikan laporan dengan format tampilan dan sumber data yang kaya dan beragam.

Pentaho Reporting ini terdiri dari 2 aplikasi yaitu :

- *Reporting Engine* yang dapat dimasukkan sebagai bagian dari aplikasi Java (*embeddable*).

- *Report Designer* yang merupakan utilitas grafik untuk merancang dan menjalankan laporan.

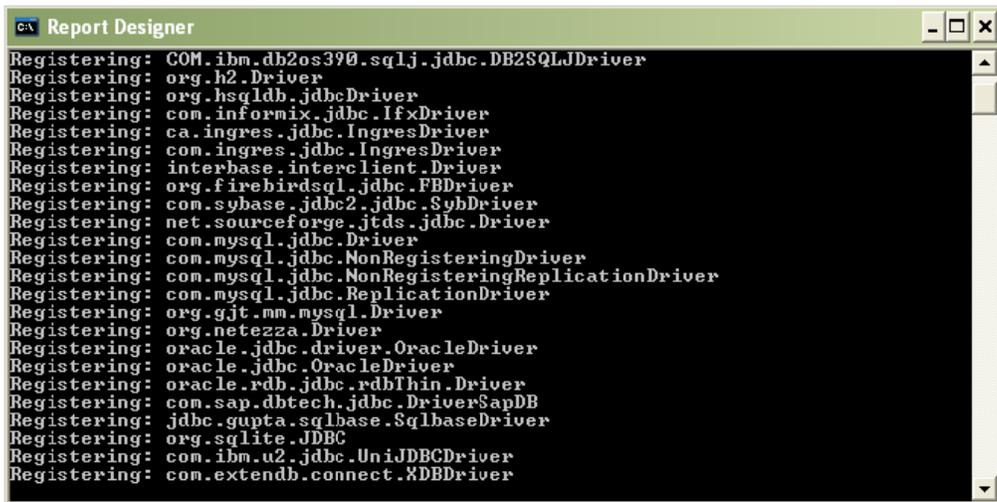
Adapun fitur-fitur dari Pentaho Reporting ini antara lain :

- Ad Hoc Report Designer Pentaho.
- Engine reporting yang bisa di-embed di aplikasi Java.
- Multi Platform.
- Dapat mengambil source dari berbagai relational engine. Lihat link berikut untuk database relasional yang didukung oleh Pentaho Reporting.
- Dapat mengambil source OLAP yang didukung driver *OLAP4J* seperti Mondrian, SQL Server Analysis Service, PALO Molap, SAP BW, dan lainnya.
- Memiliki kemampuan mengambil source Pentaho Data Integration step. Ini menjadikan source yang dapat diambil hampir menjadi tidak terbatas.
- Dilengkapi designer yang intuitif dan koleksi chart yang lengkap untuk keperluan analisa.
- Output dalam format PDF, Excel, HTML dan CSV.

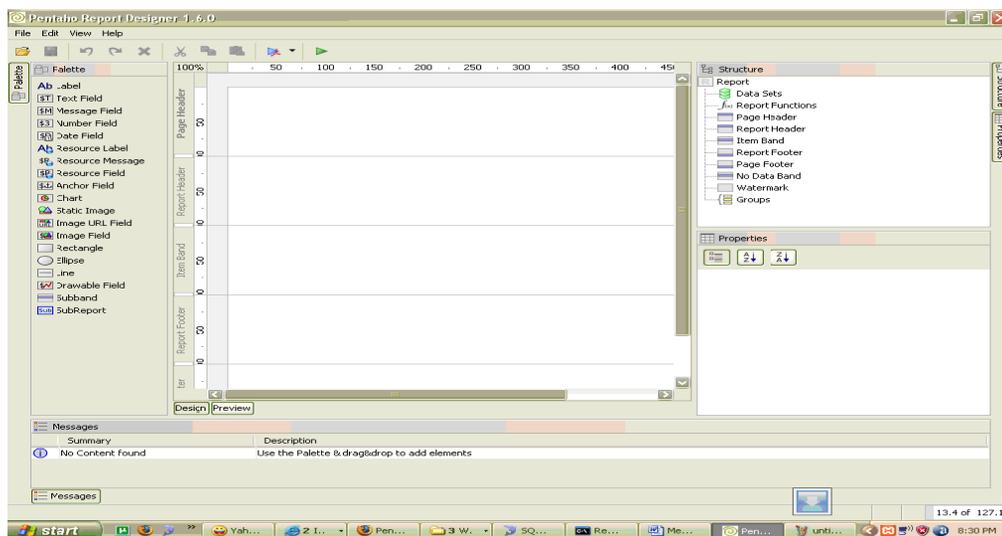
a. Pentaho Report Designer

Pentaho Report Designer adalah sebuah tools pendukung yang terdapat di dalam Pentaho suite. Fungsi utamanya adalah sebagai tools untuk membangun sebuah report/laporan yang lebih sederhana dan prosesnya lebih dapat mengefesienkan waktu dalam pembuatan sebuah data laporan. Untuk dapat menggunakan pentaho report designer kita lakukan sebagai berikut :

1. Jalankan Pentaho report designer yang telah di install, dengan cara mengklik Start → Pentaho → Pentaho Open BI suite → Report Designer.
2. Jika proses berjalan ditunjukkan oleh gambar berikut :



Gambar 2 : prompt service dari Report Designer



Gambar 3 : ruang kerja Report Designer

3.2 Mondrian / Pentaho Analysis Server

Mondrian merupakan OLAP server yang berjalan di atas platform Java dan bersifat free opensource dan telah digunakan di berbagai framework BI seperti Pentaho dan

Jasper.Inisiator Mondrian adalah Julian Hyde (<http://julianhyde.blogspot.com/>) dan sejak diakuisisioleh Pentaho tetap merupakan project leader dari Mondrian. Mondrian juga dikenal dengannama Pentaho Analysis


```

foreignKey="sk_karyawan"><Hierarchy
hasAll="true" allMemberName="Semua
Karyawan"
primaryKey="sk_karyawan" defaultMember="Se
mua Karyawan"><Table name="hrd_dim_karya
wan"/><Level name="Karyawan"
column="sk_karyawan"
nameColumn="nm_pgw" uniqueMembers="false
"/></Hierarchy></Dimension>
....
</Cube>
f. Virtual Cube
<Cube name="Sales" defaultMeasure="Unit
Sales">...<CalculatedMember name="Profit"
dimension="Measures"><Formula>[Measures].[
Store Sales] - [Measures].[Store
Cost]</Formula>...</CalculatedMember></Cub
e><VirtualCube name="Warehouse and Sales"
defaultMeasure="Profit"
>...<VirtualCubeMeasure
cubeName="Sales" name="[Measures].[Profit]"
/></VirtualCube>
g. Role
<Role name="California
manager"><SchemaGrant
access="none"><CubeGrant cube="Sales"
access="all"><HierarchyGrant
hierarchy="[Store]"
access="custom" topLevel="[Store].[Store
Country]"><MemberGrant
member="[Store].[USA].[CA]"
access="all" /><MemberGrant
member="[Store].[USA].[CA].[Los
Angeles]" access="none"
/></HierarchyGrant><HierarchyGrant
hierarchy="[Customers]"
access="custom" topLevel="[Customers].[State
Province]" bottomLevel="[Customers].[City]"><M
emberGrant member="[Customers].[USA].[CA]"
access="all" /><MemberGrant
member="[Customers].[USA].[CA].[Los
Angeles]" access="none"
/></HierarchyGrant><HierarchyGrant
hierarchy="[Gender]" access="none"
/></CubeGrant></SchemaGrant></Role>

```

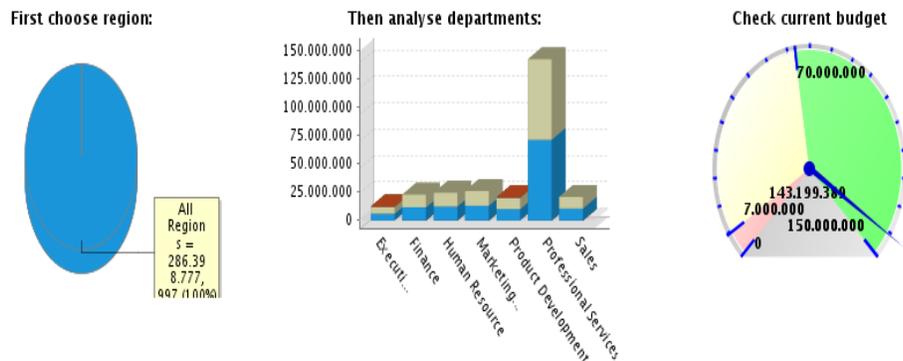
```

h. Properties
<Level name="MyLevel"
column="LevelColumn"
uniqueMembers="true"><Property
name="MyProp"
column="PropColumn"/></Level/>

```

3.3 Pentaho Dashboard

Dashboard diharapkan dapat membantu pemilik perusahaan dalam melakukan manajemen inventori. Informasi yang disajikan dalam bentuk dashboard ini diharapkan dapat membantu pemilik perusahaan dalam pengambilan keputusan. Pembuatan dashboard dapat menggunakan berbagai macam teknologi business intelligence. Seperti oracle, mysql dan pentaho. Macam Dashboard dibedakan menjadi 3. Dashboard tactical, strategic dan eksekutif. Perbedaan antara ketiganya adalah dari segi pemanfaatannya. Dashboard tactical lebih ke arah operasional, dashboard strategic menghubungkan antara operasional dengan visi misi perusahaan, apakah aktivitas perusahaan sudah sejalan dengan visi misi perusahaan. Dashboard eksekutif memantau aktivitas perusahaan untuk jangka waktu panjang. Dashboard digunakan untuk menampilkan variable-variable tertentu yang merupakan KPI (key performance index), sehingga perlu dipantau dari waktu ke waktu. Berikut adalah contoh dashboard yang bisa anda akses dari BI Developers Examples > CDF > samples > charts sample > Chart samples



Gambar 5 : Contoh Dashboard

3.4 Pentaho Data Integration (ETL)

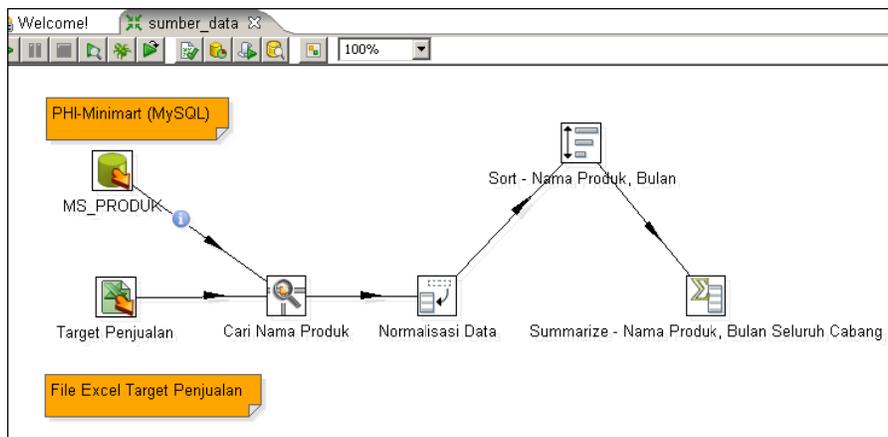
Pentaho memiliki berbagai macam sumber data seperti :

- Relasional dengan memasukkan perintah SQL (menggunakan driver tiap relasional database).
- OLAP Engine dengan perintah [MDX](#) (menggunakan driver Mondrian atau OLAP4J).
- Pentaho Metadata dengan menggunakan perintah [MQL](#).
- Pentaho Data Integration (ETL) dengan memasukkan nama step yang digunakan.

Dan Pentaho Reporting tentunya memiliki kemampuan untuk mengambil data dari salah satu sumber data tersebut. Pada artikel kali ini akan ditunjukkan penggunaan data source ETL Pentaho Data Integration 4.1 pada Pentaho Reporting 3.8.

a. Rancangan Transformasi ETL

Berikut adalah rancangan transformasi ETL yang mengambil data dari table "ms_produk" database PHI-Minimart (MySQL) dan file excel Target Penjualan.xls. Dengan demikian mengambil dua sumber data yang berbeda.



Gambar 6 : Rancangan Transformasi ETL

Transformasi tersebut mengambil data dari file Excel untuk kemudian mengambil referensi nama produk yang ada pada table MySQL karena pada file Excel tersebut tidak memiliki data nama produk.

Setelah itu data dinormalisasi, diurutkan berdasarkan nama produk & bulan dan

terakhir dihasilkan data summary berdasarkan nama produk dan bulan.

Data preview dari step "Summarize - Nama Produk, Bulan Seluruh Cabang" tampak pada gambar berikut di bawah ini. Nah, data inilah yang akan kita sajikan pada Pentaho Report. (File transformation ini dapat Anda download pada bagian akhir artikel ini).

#	nama_produk	Bulan	Target Penjualan
1	air mineral 600 ml	1	14693172
2	air mineral 600 ml	10	11581730
3	air mineral 600 ml	11	14718841
4	air mineral 600 ml	12	14745954
5	air mineral 600 ml	2	11941432
6	air mineral 600 ml	3	12952586
7	air mineral 600 ml	4	12429644
8	air mineral 600 ml	5	12535661
9	air mineral 600 ml	6	14565451
10	air mineral 600 ml	7	12919453
11	air mineral 600 ml	8	14796763
12	air mineral 600 ml	9	13237422
13	alpukat 1 kg	1	70618915
14	alpukat 1 kg	10	71134679
15	alpukat 1 kg	11	72968059
16	alpukat 1 kg	12	68179664
17	alpukat 1 kg	2	58272822
18	alpukat 1 kg	3	74642903
19	alpukat 1 kg	4	70777713

Gambar 6 : Contoh Data Preview

3.5 Pentaho Data Mining

Weka terdiri dari koleksi algoritma *machine learning* yang dapat digunakan untuk melakukan generalisasi / formulasi dari sekumpulan *data sampling*. Walaupun kekuatan Weka terletak pada algoritma yang makin lengkap dan canggih, kesuksesan data mining tetap terletak pada faktor pengetahuan manusia implementornya. Tugas pengumpulan data yang berkualitas tinggi dan pengetahuan pemodelan dan penggunaan algoritma yang tepat diperlukan untuk menjamin keakuratan formulasi yang diharapkan.

a. Algoritma Pemodelan / Classifier

Weka saat ini sudah cukup banyak mendukung algoritma untuk pemodelan data atau biasa disebut *classifier*, diantaranya adalah sebagai berikut :

- *J48*, atau kloning versi open source dari algoritma *C4.5* yang dapat digunakan untuk pembentukan pohon keputusan (*decision tree*)
- *Linear Regression*, algoritma untuk menghasilkan formulasi numerik dengan metode statistik regresi linear
- *Naive Bayes*, salah satu classifier numerik
- dan lain-lain

b. Format Data Masukan / Input

Format-format data yang dapat dijadikan input Weka adalah sebagai berikut :

- *Attribute-Relation File Format (ARFF)*, adalah tipe file teks yang berisi berbagai *instance* data yang berhubungan dengan suatu set atribut data yang dideskripsikan juga dalam file tersebut.

iris.arff - contoh data Weka dengan format ARFF

```
@RELATION iris
@ATTRIBUTE sepallength REAL
@ATTRIBUTE sepalwidth REAL
@ATTRIBUTE petallength REAL
@ATTRIBUTE petalwidth REAL
@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}
@DATA
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
```

```
4.7,3.2,1.3,0.2,Iris-setosa
4.6,3.1,1.5,0.2,Iris-setosa
5.0,3.6,1.4,0.2,Iris-setosa
5.4,3.9,1.7,0.4,Iris-setosa
4.6,3.4,1.4,0.3,Iris-setosa
```

- *Comma Separated Values (CSV)*, file teks dengan pemisah tanda koma (,) yang cukup umum digunakan *data.csv*

```
Nama, Kondisi, Aksi
Feris, lapar, makan
Feris, kenyang, tidur
Budi, lapar, makan
Budi, kenyang, tidur
PHI, lapar, makan
PHI, kenyang, makan
Jon, lapar, makan
Jon, kenyang, tidur
```

- *Format C4.5*, untuk penjelasan dan contoh dari format file ini dapat dilihat di <http://www.cs.washington.edu/dm/vfml/apendixes/c45.htm>
- *Serialisasi* data biner oleh Weka

4. KESIMPULAN DAN SARAN

a. Kesimpulan

Pentaho adalah kumpulan aplikasi *Business Intelligence* (BI), Pentaho telah didopsi untuk melakukan pengolahan data besar oleh berbagai institusi/perusahaan di Indonesia dan sampai saat ini sudah cukup banyak. Pentaho memiliki berbagai fungsi dengan produk yang dimiliki antara lain Pentaho, Pentaho Analysis, Pentaho Dashboards, Pentaho Data Integration (ETL), Pentaho Data Mining.

b. Saran

Untuk mengerti dan memahami Aplikasi/ Tool Pentaho bahasa Query Sql, OLAP, MOLAP dan ROLAP. Dengan Pentaho dapat menyelesaikan integrasi data yang besar dan kompleks.

5. DAFTAR PUSTAKA

<http://pentaho.ph-integration.com/reporting/pentaho-report-designer>, diunduh 10 April 2012.

Roldal, M.C, "Pentaho 3.2 Data Integration Beginner's Guide". Packet

<http://pentaho.ph-integration.com/reporting/mengambil-berbagai-sumber-data-dari-etl-transformation>. Diunduh 10 April 2012

http://pentaho.wikia.com/wiki/Kategori:Solusi_masalah_kemiskinan_negara_indonesia.d
iunduh 10 April 2012