



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB II

LANDASAN TEORI

2.1 Gempa Bumi

Gempa bumi adalah suatu sentakan asli yang terjadi di bumi, bersumber dari dalam bumi yang kemudian merambat ke permukaan bumi, gempa bumi berkaitan dengan serangkaian gerak gelombang yang merambat dan menembus batuan penyusun bumi (Katili, 1986). Gempa bumi dapat dikategorikan kedalam beberapa jenis menurut penyebabnya (Hartuti, 2009) :

a. Gempa Tektonik

Gempa bumi yang sering terjadi di Indonesia disebabkan oleh gejala tektonik, yaitu gerakan lempeng tektonik pada lapisan kulit Bumi. Lempeng tektonik merupakan bagian dari litosfer yang padat dan terapung di atas lapisan selubung bergerak satu sama lain. Gempa ini terjadi karena pelepasan tenaga yang dihasilkan oleh pergeseran lempeng tektonik.

Jika dua lempeng bertemu pada satu sesar (patahan), kadang dapat bergerak saling menjauhi, mendekati, atau saling bergeser. Selanjutnya, terjadi pengumpulan energi yang berlangsung terus sampai pada suatu saat batuan pada lempeng tektonik tidak lagi kuat menahan gerakan tersebut. Akibatnya, terjadi pelepasan secara tiba-tiba hingga dapat menggetarkan kulit Bumi dengan kekuatan besar yang kita kenal sebagai gempa bumi tektonik.

b. Gempa Vulkanik

Gempa yang mengguncang Bumi juga dapat ditimbulkan oleh gejala vulkanik atau gunung api. Letusan gunung api yang terjadi disebabkan oleh aliran magma dari dalam Bumi menerobos ke atas lapisan kerak Bumi. Letusan gunung berapi yang keras menyebabkan getaran kulit Bumi, terutama di daerah sekeliling gunung berapi. Pengaruh gempa vulkanik tidak sampai radius jarak yang jauh. Intensitas gempa biasanya lemah sampai sedang. Akibat yang ditimbulkan oleh gempa vulkanik juga tidak sebesar gempa tektonik.

c. Gempa Runtuhan

Selain gempa tektonik dan vulkanik, gempa bumi dapat terjadi karena runtuhnya lapisan. Kegiatan penambangan bawah tanah menyisakan rongga-rongga di bawah tanah berupa gua. Apabila runtuh, permukaan Bumi akan bergetar.

Gempa jenis ini bersifat lokal dan kekuatannya paling lemah. Gempa yang mengguncang permukaan Bumi getarannya dapat dirasakan dalam radius jarak yang jauh. Ini semua karena gempa menciptakan sebuah gelombang yang disebut gelombang seismik (gelombang gempa). Gelombang seismik ini merambat ke segala arah dari sumber atau titik asal gempa di bawah tanah. Gelombang seismik ada yang merambat melalui bagian dalam Bumi dan ada yang merambat sepanjang permukaannya.

2.2 Visualisasi Data

Visualisasi data merupakan salah satu cara dalam menganalisa data dan menyajikan data tersebut ke para penggunanya. Visualisasi data mengubah bentuk data dari yang tidak dapat dimengerti oleh penggunanya menjadi sebuah bentuk grafik yang dapat dibaca dan dianalisa. Perubahan bentuk data menjadi grafik didapatkan dari pengolahan data sedemikian rupa dari rangkaian proses yang disebut data processing. Definisi visualisasi adalah menggunakan teknologi komputer sebagai pendukung untuk melakukan penggambaran data visual yang interaktif untuk memperkuat pengamatan (Card, 1998). Visualisasi dapat diinterpretasikan dalam bentuk gambar digital, GIS (*Geographic Information System*), grafik, *virtual reality*, presentasi dimensional, *graphic user interface*, *video* dan animasi (Turban, 2007).

Berdasarkan definisi dari para ahli diatas maka dapat disimpulkan bahwa visualisasi data adalah penggunaan teknologi komputer sebagai alat yang digunakan untuk menggambarkan data menjadi bentuk visual yang dilakukan setelah data diolah dan diproses sedemikian rupa untuk dapat menjadi bentuk visual yang interaktif dan dapat digunakan untuk memperkuat analisa para penggunanya. Bentuk yang dimaksud dapat berupa gambar digital, grafik ataupun video dan animasi lainnya

Tujuan utama dari visualisasi data adalah untuk mengkomunikasikan informasi secara jelas dan efektif dengan cara grafis. Bukan berarti visualisasi data harus terlihat membosankan supaya berfungsi atau sangat canggih supaya terlihat menarik. Untuk memaparkan ide secara efektif, bentuk estetis dan fungsionalitas

harus bersamaan, menyediakan wawasan bagi kumpulan data yang kompleks dan jarang dengan mengkomunikasikan aspek-aspek kunci dengan cara yang intuitif. Namun perancang terkadang gagal mencapai keseimbangan antara bentuk dan fungsi, menciptakan visualisasi data yang menawan yang gagal menyediakan tujuan utamanya untuk mengkomunikasikan informasi (Friedman, 2008).

Terdapat juga 2 tujuan utama dari data visualisasi, yakni sebagai berikut:

1. *Explain data to solve specific problems* : visualisasi dapat membantu pengguna seperti untuk mengambil keputusan, menjawab sebuah pertanyaan dan menyampaikan informasi pada suatu masalah.
2. *Explore large data sets for better understanding* : eksplorasi visual akan memberikan banyak dimensi dari kumpulan data, atau membandingkan set data dengan data lain, sehingga dapat menarik pengguna untuk mengeksplor visual tersebut, timbul pertanyaan-pertanyaan selama proses, dan menjawab setiap pertanyaan yang ada.

(Chart, 2013).

UMMN

2.3 *Data warehouse*

Data warehouse adalah suatu kumpulan data yang bersifat *subject-oriented*, *integrated*, *time-variant*, dan *non-volatile* dalam mendukung proses pengambilan keputusan (Connolly, 2002).

Data warehouse adalah sekumpulan data yang memiliki sifat subjek berorientasi, terpadu, waktu-varian, dan tetap pada pengumpulan data untuk mendukung proses pengambilan keputusan manajemen (Inmon, 2005).

Data warehouse adalah sebuah basis data komprehensif yang mendukung semua analisis keputusan yang diperlukan oleh suatu organisasi dengan menyediakan ringkasan dan rincian informasi (Turban, 2007).

Dengan definisi para ahli di atas, dapat disimpulkan bahwa *data warehouse* merupakan sebuah kumpulan data atau basis data yang mempunyai sifat subjek orientasi dan saling terhubung satu sama lainnya yang dapat digunakan untuk mendukung analisis dan menyediakan informasi rinci untuk pengambilan keputusan bagi suatu organisasi ataupun individu yang memerlukannya.

2.4 *Data Mart*

Data mart adalah bagian dari *data warehouse* yang mendukung kebutuhan dari suatu departemen atau fungsi bisnis tertentu (Connolly, 2002).

Data Mart dapat menerima data dari *data warehouse*, atau secara langsung dari sistem transaksional. *Data mart* dapat menampilkan transformasi dan kalkulasi pada data yang sama dengan *data warehouse*, tetapi *data mart* selalu terbatas pada ruang lingkungannya dan tujuan bisnis (Scheeps, 2008).

Data mart adalah *subset* dari *data warehouse* dan berfokus pada topik tertentu atau departemen, yang biasanya terdiri dari suatu subyek (Turban, 2007). Ada dua buah jenis *data mart* yang dikemukakan oleh Turban, yakni sebagai berikut:

- a. ***Dependent Data Mart***, adalah sebuah *subset* yang dibuat secara langsung dari *data warehouse*. Memiliki keuntungan dari penggunaan data model yang konsisten dan menyediakan kualitas data.
- b. ***Independent Data Mart***, adalah *data mart* yang mendukung konsep dari *single enterprise-wide data model*, tetapi *data warehouse* harus dibuat terlebih dahulu.

Dari definisi *data mart* oleh para ahli di atas, maka dapat disimpulkan bahwa *data mart* adalah sebuah bagian dari *data warehouse* yang menampilkan informasi rinci mengenai suatu informasi transaksional atau yang terdapat pada suatu departemen tertentu yang nantinya akan digunakan untuk sebagai bahan analisa dalam pengolahan data. Berbeda dengan *data warehouse*, *data mart* merupakan bagian dari *data warehouse* atau bagian kecil dari suatu departemen didalam sebuah organisasi, di dalam *data mart* biasanya berisi informasi rinci mengenai suatu transaksi proses bisnis pada suatu perusahaan atau organisasi.

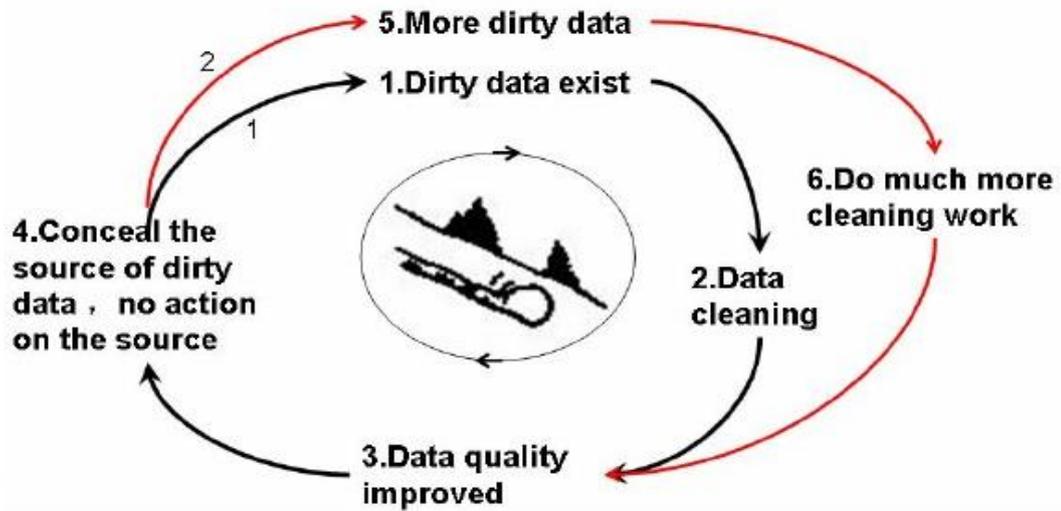
Tabel 2.1. Perbedaan *Data warehouse* dan *Data Mart*

Perbedaan	<i>Data warehouse</i>	<i>Data Mart</i>
<i>Scope</i>	Perusahaan	Departemen
<i>Subjects</i>	<i>Multiple</i>	<i>Single</i>
<i>Data Sources</i>	Banyak	Sedikit
<i>Data Sizes</i>	100 GB s.d 1 TB	Kurang dari 100 GB
<i>Implementation time</i>	1 tahun	Beberapa Bulan

2.5 *Data Cleaning*

Data cleaning merupakan salah satu tahap pada *data mining*. Data cleaning biasa disebut juga dengan *scrubbing*. Proses *data cleaning* dilakukan untuk menghilangkan kesalahan informasi pada data. *Data cleaning* dapat dilakukan dengan satu sumber atau beberapa sumber data, juga terdapat permasalahan pada level skema ataupun level *instance*. Permasalahan pada level skema dapat diselesaikan dengan perbaikan desain, *translation* dan *integration* skema. Sedangkan pada tingkat *instance* terdapat kesalahan dan tidak konsisten pada data yang merupakan fokus permasalahan yang dapat diselesaikan dengan *data cleaning* (Rahm, 2000).

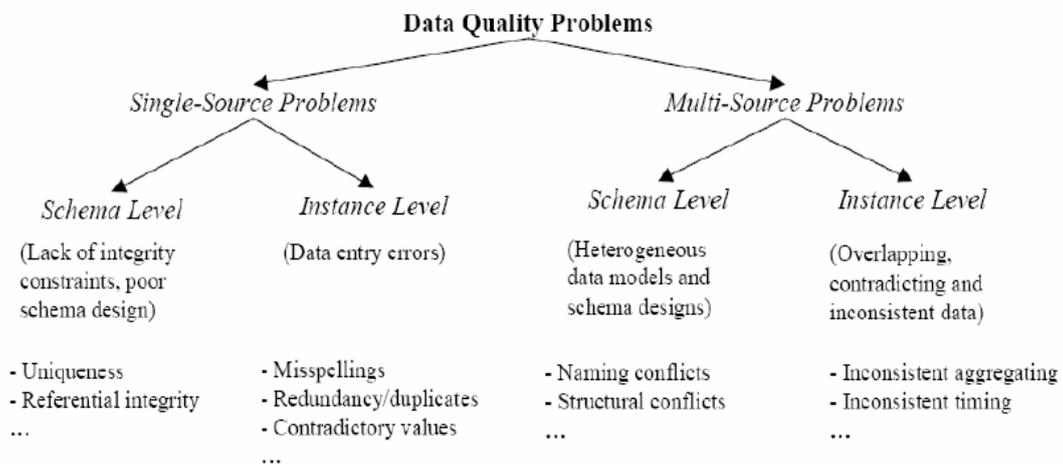
U
M
M
N



Gambar 2.1 Siklus Data Cleaning

Sumber : *Advances in Multiple Criteria Decision Making and Human*

Systems Management



Gambar 2.2 Masalah Kualitas Data

Sumber : *Data Cleaning Problems and Current Approaches*

Pendekatan pada *data cleaning* dapat dibagi menjadi beberapa fase secara umum (Rahm, 2000). Fase-fase tersebut yakni sebagai berikut :

1. *Data Analysis*

Pada fase ini dilakukan untuk mendeteksi adanya jenis kesalahan dan inkonsestensi yang akan dihapus nantinya, selain itu pemeriksaan manual dari data atau sampel data harus dilakukan untuk memperoleh sifat-sifat data dan mendeteksi masalah kualitas data.

2. *Definition of transformation workflow and mapping rules*

Data cleaning bergantung pada besarnya data, tingkat kesamaan dan tingkat kekotoran dari data. Transformasi dan *cleaning* dari banyak data mungkin harus dilakukan apabila tingkat kekotoran pada data semakin besar. Pengontrolan dan pembersihan data dapat dijalankan dengan proses yang dinamakan ETL (*Extract, Transform, Load*).

3. *Verification*

Kebenaran dan keefektifan dari transformasi atau perubahan data yang akan dilakukan harus di eveluasi dan dicoba lagi untuk menghilangkan kesalahan atau *error* dari data, sebelum transformasi data benar-benar dilakukan.

4. *Transformation*

Eksekusi dari transformasi yang telah direncanakan sebelumnya. Transformasi ini dapat juga dilakukan dengan menjalankan proses ETL pada data.

5. *Backflow of cleaned data*

Setelah data dibersihkan dengan menjalankan transformasi satu kali, ada baiknya untuk mengubah dan membersihkan data sekali lagi hingga data benar-benar bersih dari bias atau *error* dimasa depan.

2.6 DBMS (*Database Management System*)

Database Management System atau DBMS adalah suatu program komputer yang digunakan untuk memasukkan, mengubah, menghapus manipulasi dan memperoleh informasi dengan praktis dan efisien (Kadir, 2003).

Database Management System adalah sebuah sistem *software* yang memungkinkan pengguna untuk mendefinisikan, membuat, memelihara, dan mengontrol akses ke *database* (Connolly, 2002).

Beberapa keuntungan dalam pemakaian DBMS :

1. Mengontrol pengulangan data.
2. Data yang konsisten.
3. Meningkatkan keamanan.
4. Pemakaian data secara bersama-sama.
5. Meningkatkan integritas data.

Selain itu, DBMS juga mempunyai beberapa kekurangan :

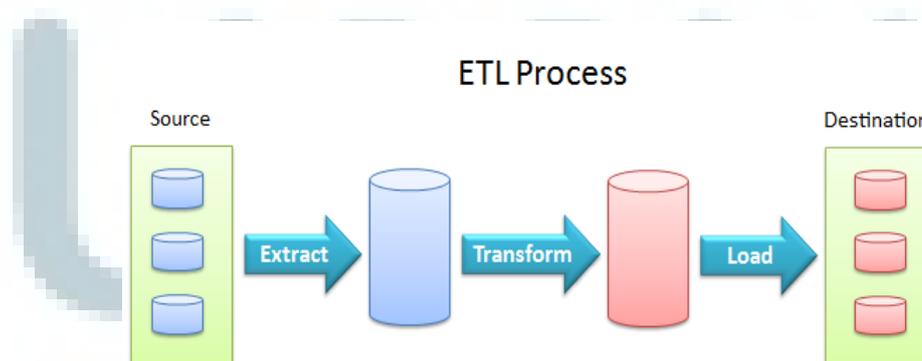
1. Kompleksitas.
2. Ukuran yang besar karena kompleksitas yang terjadi.
3. Tambahan biaya untuk *hardware*.
4. Biaya konversi yang dibutuhkan untuk penggunaan DBMS yang baru.
5. Peforma kinerja menurun.

DBMS mempunyai beberapa fasilitas yang mendukung pengolahan dan penyimpanan data (Connolly, 2002), yakni sebagai berikut :

1. DDL (*Data Definition Language*) adalah suatu bahasa yang memperbolehkan DBA (*Database Administrator*) ataupun *user* untuk menentukan tipe, struktur, dan batasan keamanan data yang akan disimpan pada *database*.
2. DML (*Data Manipulation Language*) adalah suatu bahasa yang memiliki fasilitas untuk mengoperasikan data yang ada dalam *database*.
3. SQL (*Structured Query Language*) adalah suatu bahasa yang memberikan layanan akses terhadap data.

2.7 ETL (*Extract, Transform, Load*)

ETL adalah suatu proses mengambil dan mengirim data dari sumber data ke *data warehouse*. Dalam proses pengambilan data, data harus bersih agar didapat kualitas data yang baik. Contohnya ada nomor telepon yang *invalid*, ada kode buku yang tidak eksis lagi, ada beberapa data yang *null*, dan lain sebagainya. Pendekatan tradisional pada proses ETL mengambil data dari data sumber, meletakkan pada *staging area*, dan kemudian mentransformasi dan meng-*load* ke *data warehouse* (Rainardi, 2008).



Gambar 2.3 Alur proses ETL

Sumber : Definisi *Data Mart*, *Data warehouse* dan ETL.

Terdapat 3 proses utama yang dilakukan dalam proses ETL (Rainardi, 2008), yakni sebagai berikut :

1. *Extract*

Extract atau ekstraksi adalah proses dimana data diambil atau diekstrak dari berbagai sistem operasional, baik menggunakan *query*, atau aplikasi ETL seperti *Pentaho*. Pada proses ini tabel-tabel dan informasi yang akan digunakan ditentukan dan diambil untuk nantinya akan dijadikan menjadi sumber data.

2. *Transform*

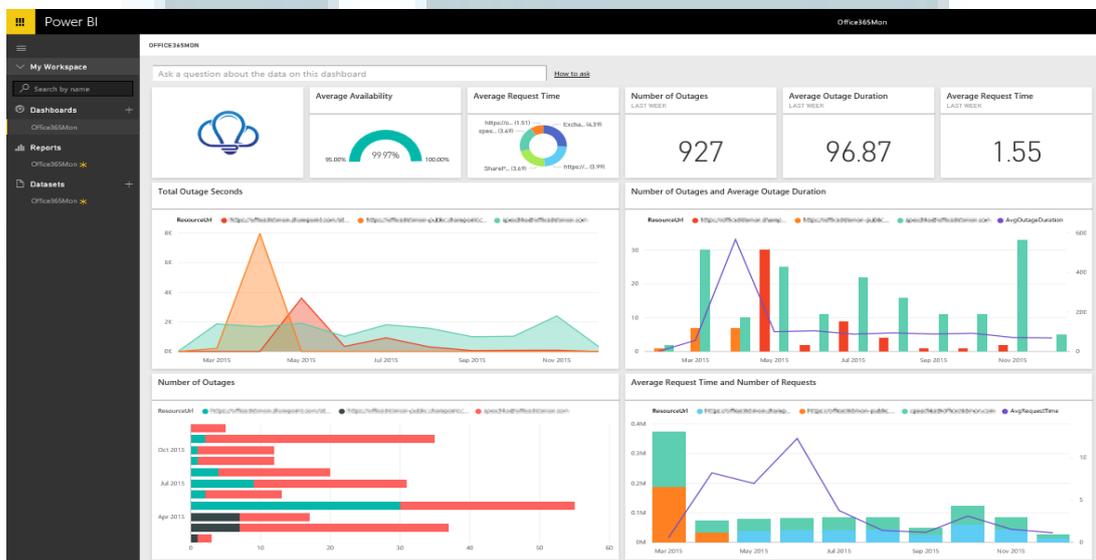
Transform atau transformasi data adalah proses dimana data mentah hasil ekstraksi disaring dan diubah sesuai dengan kebutuhan pengguna. Dalam tahap transformasi data ini digunakan beberapa rangkaian aturan atau fungsi untuk mengambil data dari sumber data yang selanjutnya akan dimasukkan ke dalam data yang baru.

3. *Load*

Load adalah Proses terakhir yang perlu dilakukan adalah proses pemuatan data yang didapatkan dari hasil transformasi ke dalam *data warehouse* untuk disimpan sebagai *archive*. Tahap *load* merupakan tahap terakhir dalam proses ETL, dimana pada tahap ini data yang telah dibersihkan dan diolah melalui 2 tahap sebelumnya dimasukkan ke dalam sebuah *data warehouse* yang nantinya data sudah bersih ini akan dapat diakses kembali melalui *data warehouse*.

2.8 Dashboard

Dashboard biasanya disebut *executive dashboard* atau *management cockpit* yang menyediakan akses informasi yang cepat dan tepat waktu, dan akses langsung ke laporan manajemen. *Dashboard* sangat *user-friendly* dan didukung dengan grafik. *Dashboard* memperbolehkan manajer untuk memeriksa laporan tertentu dan laporan yang rinci (Rainer, 2011).



Gambar 2.4 Dashboard Power BI

Sumber : <https://dpspowerbi.blob.core.windows.net>

Dashboard adalah satu kategori dari aplikasi *business intelligence* yang secara *real time* akan memperlihatkan berbagai informasi yang dibutuhkan oleh suatu organisasi atau perusahaan dengan berbagai macam format seperti *graphical gadgets, typically, gauges, charts, indicators, dan color-coded maps* yang memungkinkan mereka membuat keputusan pintar secara cepat (Kusnami, 2009).

2.9 *Microsoft Power BI*

Microsoft Power BI adalah sebuah *software* atau aplikasi yang dirancang oleh *Microsoft* dan ditujukan untuk setiap orang yang membutuhkan penyajian visualisasi data yang interaktif dan *modern* guna untuk melakukan suatu analisis mengenai data tersebut (Lachev, 2015). *Microsoft Power BI* sendiri mempunyai banyak sekali layanan atau fitur yang mempermudah penggunaanya dalam membuat sebuah grafik ataupun skema dalam memperlihatkan informasi masa lalu, masa kini, dan *trend* di masa depan.



Gambar 2.5 Logo *Microsoft Power BI*

Sumber : <https://powerbi.Microsoft.com/>

Dalam membuat sebuah *dashboard* yang informatif dan tepat bagi penggunaanya tentu membutuhkan beberapa prinsip yang harus diterapkan ke dalam *dashboard* tersebut. Terdapat 6 prinsip yang harus diperhatikan saat membangun sebuah *dashboard* (Hariyanti, 2008). Yakni sebagai berikut :

1. Menyajikan informasi mengenai *key performance index* dengan tujuan yang spesifik.
2. Mensinergikan informasi dari berbagai aspek dalam layar tunggal.
3. Merupakan alat yang responsif dan interaktif dengan penggunaanya.

4. Memungkinkan tiga hal sekaligus yaitu analisis kondisi sebelumnya, memperlihatkan kondisi saat ini, dan memprediksi tren kedepan.
5. Memiliki faktor personalisasi. Setiap bagian dalam organisasi memiliki *dashboard* sendiri.
6. Memungkinkan kolaborasi dan komunikasi antar bagian dalam organisasi.

Beberapa prinsip diatas dapat dipenuhi oleh *Microsoft Power BI* sebagai *tools* dalam visualisasi data dan beberapa prinsip diantaranya tergantung pada pembuat *dashboard* dalam menerapkan prinsip-prinsip tersebut.

2.10 *Open Refine*



Gambar 2.6 Logo *Open Refine*

Sumber : <http://openrefine.org/>

Open Refine adalah salah satu *software data cleaning* yang dibuat oleh *Google* dan dijalankan pada sistem operasi *windows*. *Open Refine* digunakan dalam mengakses data dari jumlah kecil hingga besar, membersihkan dan merubah data, serta melakukan rekonsiliasi dan perbandingan data (Verborgh, 2013).

Dalam Penelitian ini, *Open Refine* akan digunakan sebagai *software* pembersih data atau *data cleaning*, dimana pembersihan data seperti penghapusan kolom, perubahan nama kolom, perubahan jenis data, penghapusan data, perubahan data dilakukan didalam *Open Refine*.

2.11 *Pentaho Data Integration*



Gambar 2.7 Logo *Pentaho Data Integration*

Sumber : <http://www.bizcubed.com.au>

Pentaho merupakan sebuah *software* yang dirancang oleh perusahaan bernama *Hitachi Group*. *Software* ini dapat melakukan banyak proses integrasi data secara menyeluruh termasuk melakukan proses ETL.

Pentaho digunakan didalam penelitian ini untuk merubah atau mentransformasi data dari bentuk *Excel* menjadi bentuk *Mysql* yang disimpan didalam *database*. Selain itu, *Pentaho* juga digunakan untuk memverifikasi atau validasi data yang telah dibersihkan (Casters, 2010), hal ini dilakukan untuk memastikan bahwa data yang telah disimpan di *database* telah benar-benar bersih dari bias atau *error* .

U M N

2.12 *Visual Data Mining (VDM)*

Visual Data Mining (VDM) merupakan sebuah panduan dalam mengerjakan atau membangun sebuah visualisasi data. Dalam metode ini terdapat 3 fase dan 8 tahap dalam mengerjakan visualisasi data (Soukup, 2002). 3 fase dan 8 tahap tersebut yakni sebagai berikut :

A. *Project Planning Phase*

Fase ini merupakan fase dimana penelitian direncanakan apa saja yang dibutuhkan dalam mengerjakannya. Pada fase ini terdiri dari 2 tahap, yakni sebagai berikut :

1. *Justify and Plan the Project*

Tujuan pada tahap ini adalah untuk merencanakan *project* visualisasi yang akan dibuat seperti tujuan, hasil, dan kebutuhan dari visualisasi

2. *Identify the top Business Question*

Tahap ini mengharuskan untuk melakukan wawancara kepada narasumber atau pengguna visualisasi data, dengan tujuan untuk dapat mengetahui kebutuhan visualisasi yang akan dibuat

B. *Data Preparation Phase*

Pada fase ini penulis diharuskan untuk menyiapkan segala kebutuhan yang telah direncanakan pada tahap sebelumnya. Fase ini terdiri dari 3 tahap, yakni sebagai berikut :

1. *Choose the Data Set*

Setelah data dikumpulkan, maka pada tahap ini dilakukannya pemilihan data yang sesuai dengan dengan kebutuhan visualisasi nantinya.

2. *Transform the Data Set*

Pada tahap ini dilakukannya perubahan bentuk data, dimana data diolah sedemikian rupa untuk terbebas dari *error*. Proses Data Cleaning dan ETL dilakukan pada tahap ini.

3. *Verify the Data Set*

Setelah data dibersihkan dan dianggap bersih, maka tahap selanjutnya adalah memastikan kembali kebersihan dari data yang dianggap bersih tersebut dengan melakukan proses validasi atau verifikasi.

C. *Data Analysis Phase*

Pada fase yang terakhir ini, penulis bertugas untuk memproses dan menganalisa data yang telah direncanakan dan disiapkan sebelumnya. Fase ini terdiri dari 3 tahap, yakni sebagai berikut :

1. *Choose the Visualization for Mining Tool*

Pada tahap ini dilakukannya pemilihan *tools* yang akan digunakan dalam visualisasi, dalam penelitian ini akan menggunakan *tool* bernama *microsoft power BI* sebagai *visualization tools*.

2. *Analyze the Visualization or Mining Model*

Setelah visualisasi selesai dibangun, maka tahap selanjutnya adalah melakukan analisis terhadap visualisasi tersebut. Hal ini dilakukan untuk mendapatkan informasi yang terdapat pada visualisasi.

3. *Verify and Present the Visualization or Mining Model*

Setelah analisa selesai dilakukan maka tahap selanjutnya adalah mengverifikasi hasil analisa dan visualisasi kepada pengguna atau narasumber dari data yang digunakan.

UMMN