



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 Teknologi

Harahap menjelaskan bahwa penggunaan kata teknologi pada dasarnya mengacu pada sebuah ilmu pengetahuan yang menyelidiki tentang cara kerja di dalam bidang teknik, serta mengacu pula pada ilmu pengetahuan yang digunakan dalam pabrik atau industry tertentu. Definisi ini tentu saja sangat mengacu pada definisi praktis dari teknologi, yang banyak ditemukan pada pabrik-pabrik dan juga industry tertentu. (Harahap, 1982)

“Teknologi sebagai suatu cara praktis yang menjelaskan mengenai cara kita semua sebagai manusia membuat segala sesuatu yang berada di sekita sini.” Pengertian ini merujuk pada penggunaan teknologi yang merupakan seluruh benda yang dibuat oleh manusia, dimana setiap orang bisa saja membuat dan juga mengembangkannya apabila mempelajarinya dengan baik dan dapat menerapkannya secara praktis. (Franklin 1989)

2.2 Data Visualization

Data visualisasi merupakan sebuah teknik yang digunakan untuk mengomunikasikan data atau informasi dengan cara menggunakan menerjemahkannya menjadi bentuk sebagai objek visual (contoh. *Points*, garis, atau *bar*) yang memiliki grafik. Tujuan dari data visualisasi ini sendiri ialah untuk mengomunikasikan informasi dengan jelas serta secara efisien untuk pengguna.

Data visualisasi merupakan salah satu langkah dalam data analisis serta *data science*.

Menurut Friedman (2008) tujuan utama dari data visualisasi adalah untuk mengomunikasikan informasi secara jelas dan efektif melalui pendekatan grafis. Data visualisasi tidak selalu harus terlihat membosankan untuk dapat berfungsi atau bahkan terlihat sangat baik dan *sophisticated* untuk terlihat indah. Tujuannya adalah untuk menjelaskan ide secara efektif, baik dalam bentuk estetika ataupun dalam bentuk fungsinya harus dapat menjadi kegunaan bagi setiap pihak, dan memberikan wawasan dari bagian yang langka dan kompleks dengan cara mengomunikasikan kunci aspek dari data tersebut dengan bentuk yang lebih intuitif.

Data visualisasi memiliki relasi yang sangat dekat dengan *information graphics*, *information visualization*, *exploratory data analysis*, serta *statistical graphics*. Dalam millennium baru, data visualisasi telah menjadi area aktif dari penelitian, pembelajaran dan *development*. Menurut Post et al. (2003), data visualisasi memiliki *scientific* dan informasi visualisasi yang tergabung.

2.3 Data Mining

Data mining adalah proses semi otomatis yang menggunakan statistik, matematika, kecerdasan buatan, dan mesin untuk mengekstraksi dan mengidentifikasi informasi dan pengetahuan potensial yang tersimpan di dalam database. (Turban, 2005)

“Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data.” (Larose, 2006)

Data mining adalah gabungan keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk penanganan permasalahan dan pengambilan informasi dari database yang besar. (Larose, 2006)

2.4 Tableau Software

Tableau adalah software business intelligence yang memungkinkan semua orang untuk melakukan koneksi ke data secara mudah, kemudian memvisualisasikan dan membuat *dashboard* yang interaktif dan dapat dibagikan. (Tableau, 2016)

Tableau *Desktop* adalah aplikasi eksplorasi yang memungkinkan pengguna untuk menjawab pertanyaan mendesak secara cepat. Tableau membuat sharing data menjadi mudah, tidak peduli apapun kebutuhannya. (Tableau, 2016)

Tableau Public adalah *free service* yang memungkinkan setiap orang untuk berbagi data secara interaktif melalui web. Setiap orang bisa melakukan analisis dan visualisasi dari data yang ada. (Tableau, 2016)

2.5 Tableau Data Roles

Dimension adalah data roles yang ditetapkan pada suatu *field* oleh Tableau karena mengandung *categorical data*. *Dimension* secara umum menciptakan header. (Tableau, Dimensions and Measures, 2016)

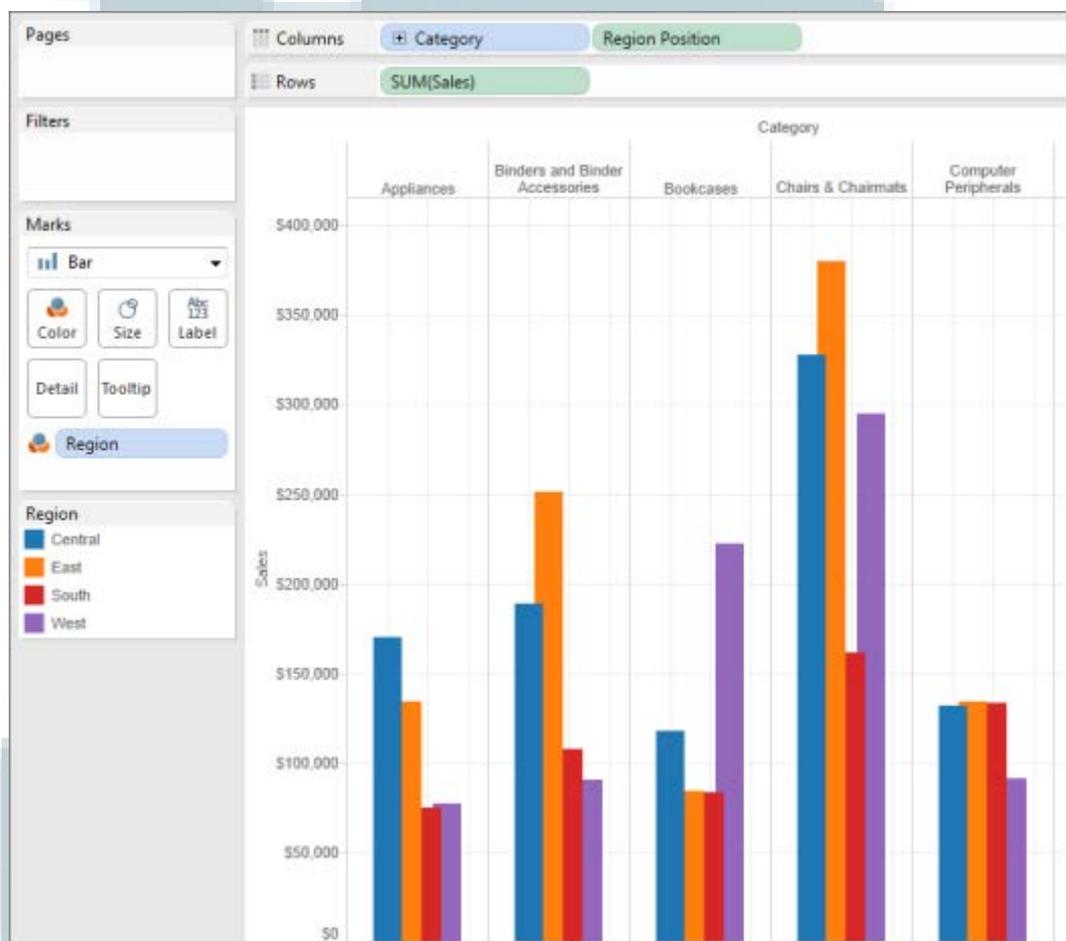
Measure adalah data roles yang ditetapkan pada suatu *field* oleh Tableau karena mengandung *quantitative data*. *Measure* secara umum menciptakan *continuous axis*. (Tableau, Dimensions and Measures, 2016)

2.6 Multidimensional Data Visualization Tools

Sebuah informasi yang terkandung di dalam data dapat lebih mudah dimengerti jika disajikan dalam perspektif yang berbeda. Banyak cara yang bisa digunakan untuk berinteraksi dengan data. Berikut adalah objek atau metode yang bisa digunakan untuk menciptakan representasi visual dari data tersebut. (Mantange, Beamon, & Huffman, 2010)

2.3.1 Multidimensional Chart

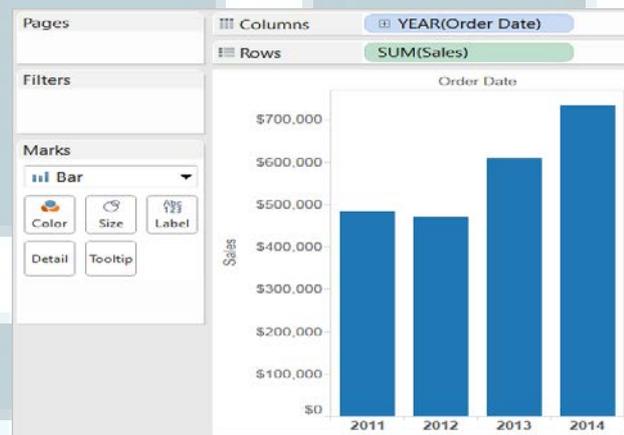
Bar chart memungkinkan *user* untuk membuat sebuah visualisasi dari dua atau enam variabel dalam satu waktu. Salah satu dari variabel bisa dilampirkan ke satu dari tiga *axes* dan satu dari tiga rentang warna. (Mantange, Beamon, & Huffman, 2010)



Gambar 2.2 Bar Chart

2.7 Building Tableau Data Views

Bar Chart digunakan untuk membandingkan data dalam kategori. *Bar Chart* dibuat dengan menempatkan *dimension* pada *rows* dan *measure* pada *columns*. (Tableau, Build a Bar Chart, 2016)



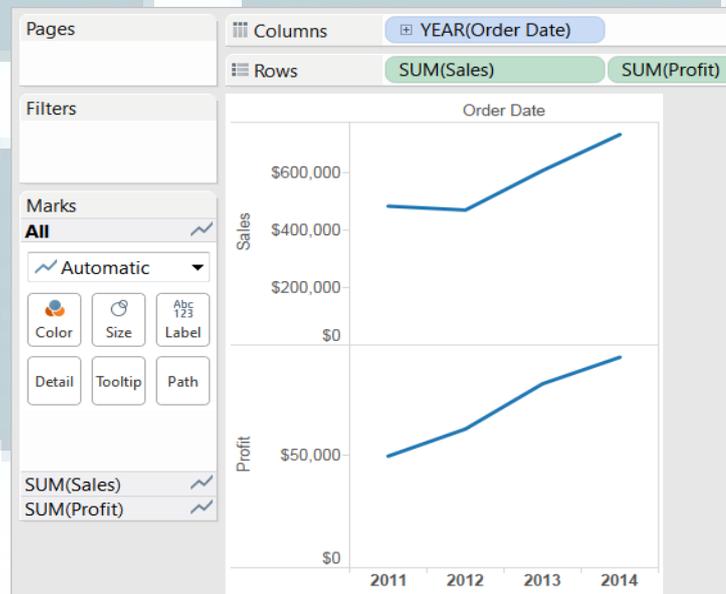
Gambar 2.5 Tableau Bar Chart

Text tables diciptakan dengan menempatkan satu *dimension* pada *rows*, *dimension* lain pada bagian *columns*, dan dilengkapi dengan menempatkan satu atau lebih *measure* pada *Marks card*. (Tableau, Build a Text Table, 2016)

Sub-Category	Order Date			
	2011	2012	2013	2014
Accessories	\$25,014	\$40,524	\$41,896	\$59,946
Appliances	\$15,314	\$23,241	\$26,050	\$42,927
Art	\$6,058	\$6,237	\$5,910	\$8,914
Binders	\$43,488	\$37,453	\$49,485	\$72,986
Bookcases	\$20,037	\$38,544	\$26,275	\$30,024
Chairs	\$77,242	\$71,735	\$83,919	\$95,554
Copiers	\$10,850	\$26,179	\$49,599	\$62,899
Envelopes	\$3,856	\$4,512	\$4,730	\$3,379
Fasteners	\$661	\$545	\$960	\$858
Furnishings	\$13,826	\$21,090	\$27,874	\$28,915
Labels	\$2,841	\$2,956	\$2,827	\$3,861
Machines	\$62,023	\$27,764	\$55,907	\$43,545
Paper	\$14,835	\$15,288	\$20,638	\$27,718
Phones	\$77,391	\$68,314	\$78,660	\$105,643
Storage	\$50,329	\$45,048	\$58,632	\$69,834
Supplies	\$14,394	\$1,952	\$14,278	\$16,049
Tables	\$46,088	\$39,150	\$60,833	\$60,894

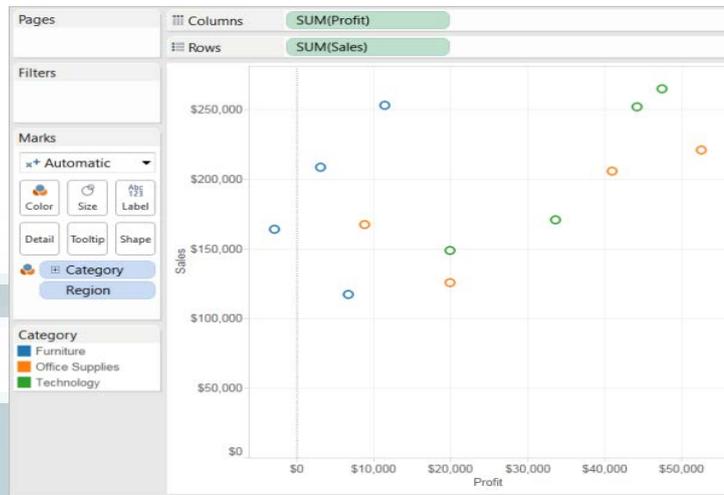
Gambar 2.6 Tableau Text Tables

Line Charts mengkoneksikan individual *data point* pada *data view*. *Line Chart* menyediakan cara yang sederhana untuk memvisualisasikan urutan atau *sequence* dari *values*. *Line Chart* sangat berguna untuk melihat *trends* dari waktu ke waktu atau untuk memprediksi *values* di masa depan. (Tableau, Building Line Charts, 2016)



Gambar 2.7 Tableau Line Charts

Scatter Plots digunakan untuk memvisualisasikan hubungan antara *numerical variables*. *Scatter Plots* dibuat dengan menempatkan setidaknya satu *measure* pada bagian *columns* dan setidaknya satu *measure* pada bagian *rows*. (Tableau, Build a Scatter Plot, 2016)



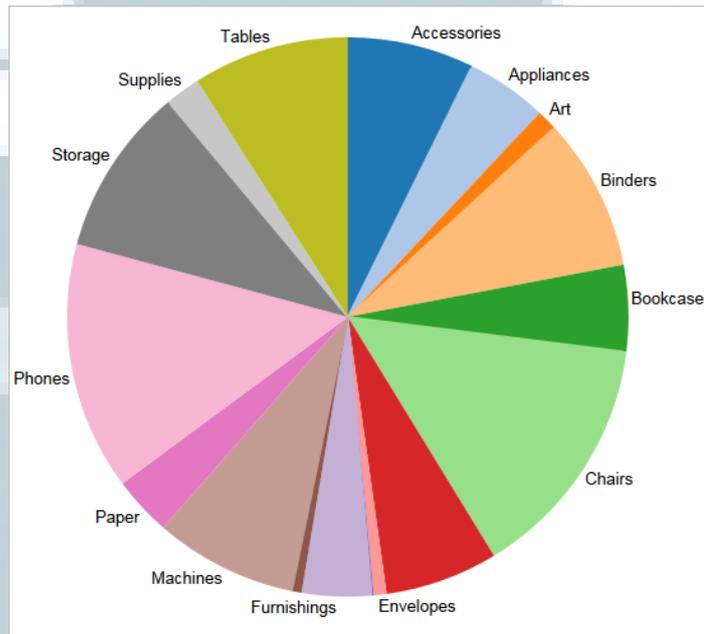
Gambar 2.8 Tableau Scatter Plot

Histogram adalah *chart* yang menampilkan *shape* atau bentuk dari sebuah distribusi. *Histogram* terlihat seperti sebuah *bar chart* tetapi mengelompokkan *value* pada *continuous measure* menjadi *ranges*. (Tableau, Build a Histogram, 2016)



Gambar 2.10 Tableau Histogram

Pie Charts digunakan untuk menunjukkan sebuah proporsi. *Pie Chart* dibuat dengan memilih *Pie mark* pada *Marks card*. (Tableau, Build a Pie Chart, 2016)



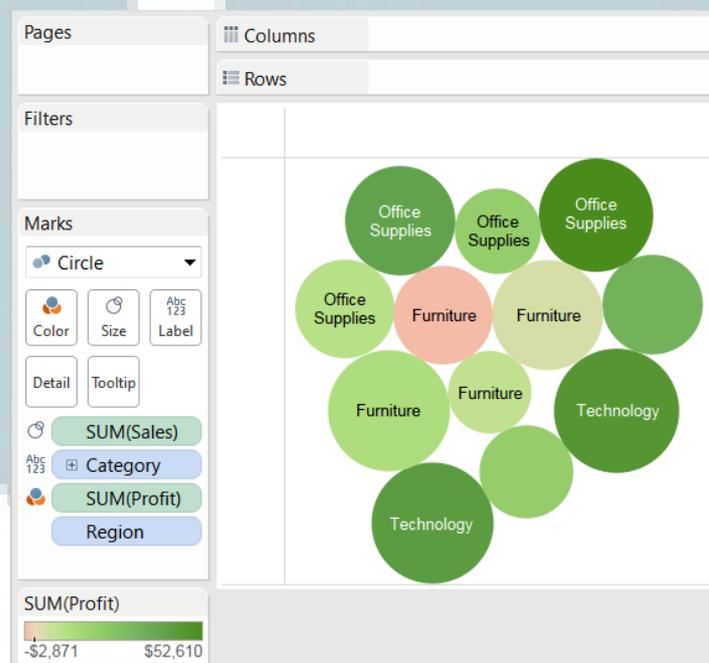
Gambar 2.12 Tableau Pie Chart

Box Plots biasanya juga dikenal sebagai *box-and-whisker plots* yang digunakan untuk menunjukkan distribusi dari *values* pada *axis*. (Tableau, Build a Box Plot, 2016)



Gambar 2.14 Tableau Box Plot

Bubble Charts digunakan untuk menampilkan data dalam sebuah *cluster* atau *circles*. *Dimension* mendefinisikan *individual bubbles*, dan *measure* mendefinisikan ukuran dan warna dari *individual circles*. (Tableau, Build a Packed Bubble Chart, 2016)



Gambar 2.15 Tableau Packed Bubble Chart

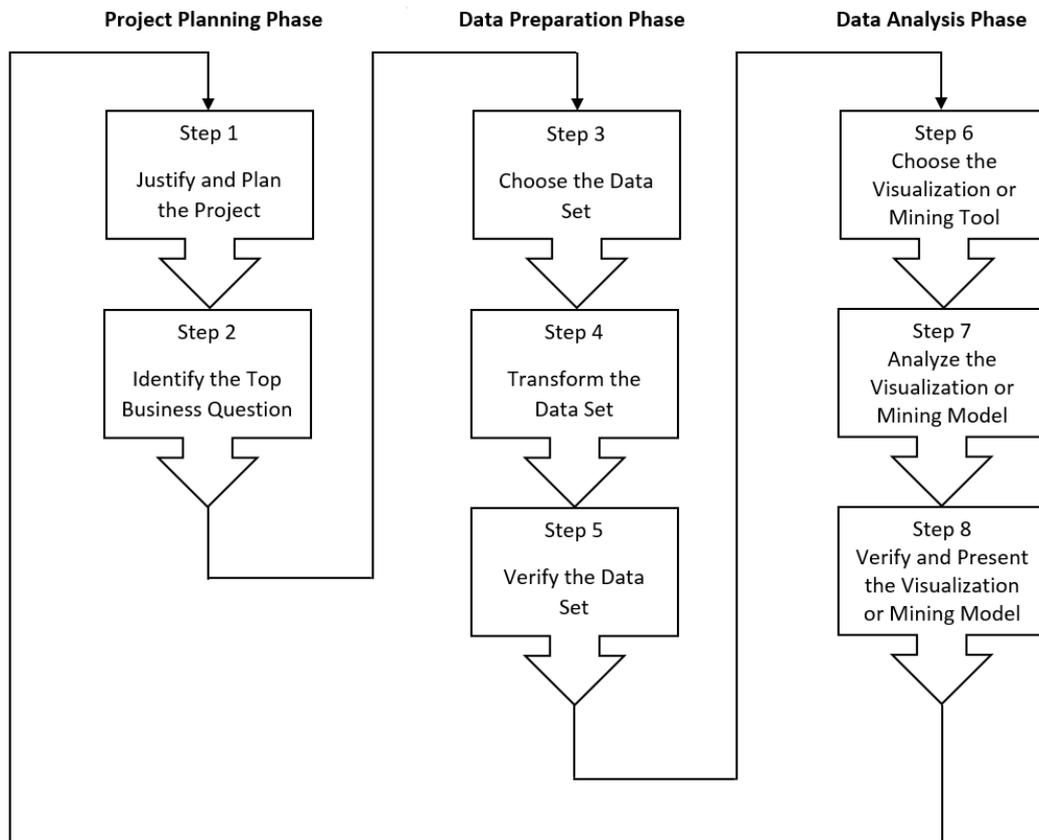
Map Views digunakan untuk menampilkan dan menganalisa *geographic data*. Untuk menciptakan *Map Views*, diperlukan *geographic fields* dan juga *measure* atau *continuous dimension* pada *Marks card*. (Tableau, Build a Map View, 2016)



Gambar 2.16 Tableau Map View

UMMN

2.8 Visualization Data Mining



Gambar 2. 1. Eight-Step Data Visualization and Visual Data Mining

Methodology

Sumber: Visual Data Mining: Techniques and Tools For Data Visualization and Mining (Soukup, 2002)

Gambar 3.4 menjelaskan bagaimana metode *Visual Data Mining* (VDM) berjalan pada penelitian ini. *Visual Data Mining Methodology* merupakan sebuah panduan siklus proyek dalam membangun sebuah data visualisasi. Pada metode ini

terdapat tahapan-tahapan mulai dari persiapan data hingga pembuatan visualisasi dari data-data tersebut. Tahapan-tahapan dalam metode ini terbagi menjadi tiga yaitu yang pertama *Project Planning Phase* yang terdiri dari *justify and plan project* dan *identify the top business question*, kedua *Data Preparation Phase* yang terdiri dari *choose the data set*, *transform the data set* dan *verify the data set*, ketiga adalah *Data Analysis Phase* yang terdiri dari *choose the visualization or mining tools*, *analyze the Visualization or Mining Model* dan *verify and present the visualization or mining model*.

2.9 Project Planning Phase (Tahap Perencanaan Proyek)

2.9.1. *Justify and Plan the Project* (Pembenaran dan rencana Proyek)

Merencanakan project dan membuat pedoman untuk estimasi waktu proyek dan kebutuhan sumber daya. Ada tiga tipe dari proyek data visualisasi yaitu

- a. *A proof-of-concept VDM project* memiliki *scope* yang terbatas.

Secara keseluruhan *scope*-nya adalah untuk menentukan apakah visualisasi dan *data mining* dapat memberikan keuntungan bagi bisnis dalam perusahaan untuk membuktikan kepada pembuat keputusan.

- b. *A pilot VDM project* juga memiliki *scope* yang terbatas. Secara keseluruhan *scope*-nya adalah untuk menginvestigasi, menganalisis, dan menjawab satu atau banyak *business question*.

c. A *production VDM project* scope-nya sama dengan *pilot project*, namun visualisasi dan model *data mining* yang dihasilkan diimplementasikan ke dalam lingkungan produksi. Secara keseluruhan *scope-nya* adalah menginvestigasi sepenuhnya, menganalisis dan menjawab satu atau banyak *business question* dan mengimplementasikannya dan mengukur hasil dari model visualisasi produksi dan *data mining* yang telah dibuat.

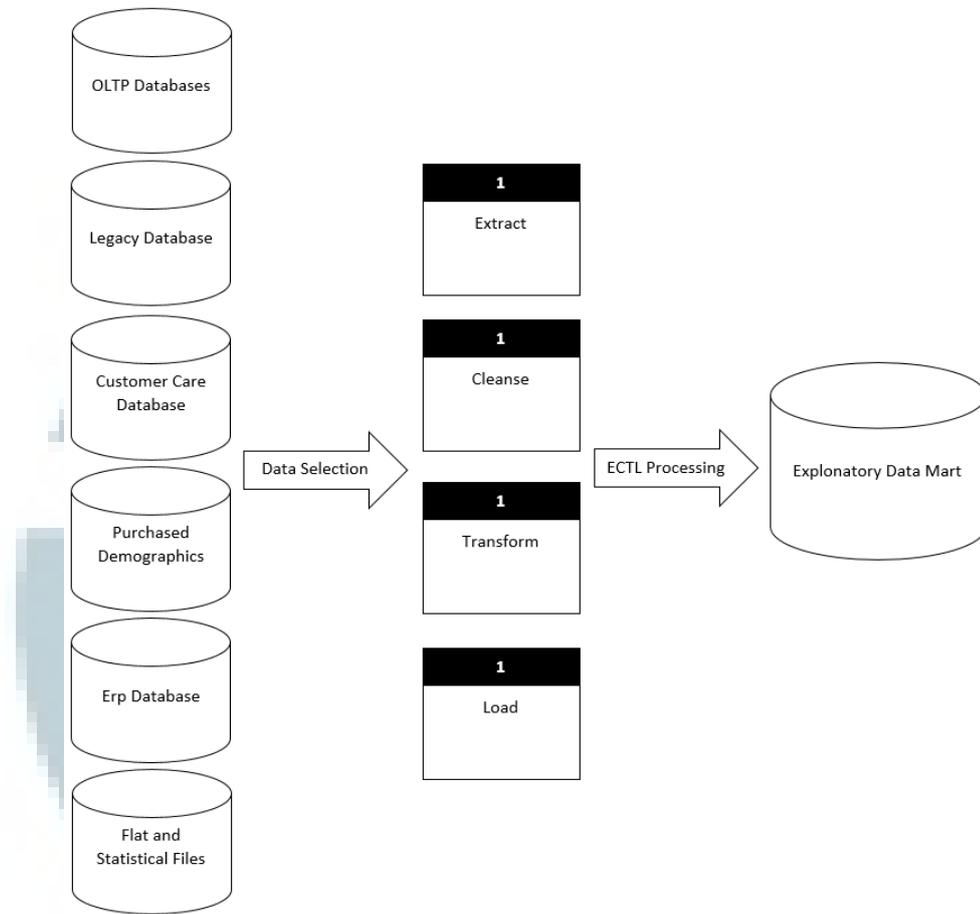
2.9.2 Identify the Top Business Question (Identifikasi Pertanyaan Bisnis)

Pada langkah ini harus memastikan *business question* apa yang butuh diinvestigasi, seperti memetakan *business question* ke dalam *problem definition* yang dapat diatasi dengan model data visualisasi dan *data mining*. Selain itu harus mencari tahu dan memperjelas apa yang sebenarnya diharapkan sebagai *output* dari proyek ini.

2.10 Data Preparation Phase (Tahap Persiapan Data)

2.10.1 Choose the Data Set (Memilih Set Data)

Berdasarkan tahap sebelumnya yaitu mengidentifikasi *top business question*, maka pada langkah ini mengharuskan untuk memilih sumber data yang dapat dianalisa dan dapat menjawab pertanyaan-pertanyaan yang ada. Dalam mendesain sebuah *exploratory data mart* terdapat tujuan utama yaitu adalah menyatukan informasi dari sumber data yang berbeda-beda ke dalam satu *exploratory data mart* dan dapat digunakan dalam mengambil keputusan.



Gambar 2. 2. Data Flow from Operational Data Source to Visualization and Data Mining Tools.

Sumber: Visual Data Mining: Techniques and Tools For Data Visualization and Mining (Soukup, 2002)

Exploratory Data Mart berisi satu buah *table* seperti yang dapat dilihat pada Gambar 2.2. untuk *proof-of-concept* dan *pilot VDM project*, mungkin hanya dibutuhkan beberapa *table* saja di dalam *exploratory data mart* sebagai contoh jika *business question* berurusan dengan *customer* dan *product revenue*,

maka hanya perlu menjalankan *exploratory* yang berisi satu atau lebih *table* yang menyangkut informasi tentang *customer* dan *product revenue* saja.

2.10.2 Transform the Data Set (Transformasi Set Data)

Pada tahap *transform the data set* terdapat dua tingkat *logical transformations* yang mungkin dibutuhkan sebelum memvisualisasikan data set yaitu: *table-level logical transformation* dan *column-level logical transformation*.

Pada *table-level logical transformation* terdapat beberapa transformasi seperti:

- a. *Transforming weighted data set* yaitu merubah atau mengganti kolom dan *record* dari *data set* dengan menambahkan baris baru yang meniru bobot aslinya.
- b. *Transformation time series data set* adalah perubahan yang menstruktur ulang *time series* dari *data set*. Beberapa data visualisasi membutuhkan *time series*.
- c. *Aggregating data set* melibatkan agregasi kumpulan data ke unit percobaan yang lebih tinggi, seperti menggabungkan data penjualan dengan Negara untuk dipresentasikan ke dalam map.
- d. *Filtering data set* melibatkan menggunakan bagian-bagian spesifik dari sekumpulan data.

Sedangkan dalam *column-level logical transformation* terdapat dua buah *group* yaitu:

- a. *Simple column transformation* merupakan transformasi kolom yang menghapus kolom, merubah tipe data kolom, atau membuat kolom baru berdasarkan *expression*. Transformasi ini dapat memungkinkan untuk memanipulasi *data set*. Seperti mengkombinasikan banyak kolom ke dalam satu kolom yang akan digunakan untuk visualisasi atau meningkatkan akurasi dari model *data mining*.
- b. *Column value grouping transformation* adalah logika transformasi yang digunakan untuk mengelompokkan nilai yang berkelanjutan dan diskrit ke dalam kelompok yang lebih besar.

2.10.3 *Verify the Data Set* (Verifikasi Set Data)

Setelah melakukan pemilihan data dan melakukan transformasi data untuk divisualisasikan, pada tahap ini tahap-tahap sebelumnya dicek kembali apakah sudah sesuai, terbebas dari error, dan tidak menimbulkan bisa yang timbul akibat proses pemilihan data maupun proses transformasi data yang ada. Untuk melakukan verifikasi keakuratan data, *data set* dibandingkan kembali dengan sumber data aslinya.

2.11 *Data Analysis Phase* (Tahap Analisis Data)

2.11.1 *Choose the Visualization or Mining Tools* (Memilih Alat Visualisasi)

Pada tahap ini membicarakan tentang bagaimana memilih *data visualization* atau *data mining tool* yang sesuai untuk *business question* yang

telah diinvestigasi. *Data visualization tool* dapat dibagi menjadi dua kategori yaitu:

- a. *Multidimensional visualization tool: Scatter, line, column, bar dan pie graph*
- b. *Specialized landscape and hierarchical visualization: map dan tree graph*

2.11.1 *Analyze the Visualization or Mining Model* (Menganalisa Model Visualisasi)

Setelah memilih data visualisasi apa yang digunakan, selanjutnya adalah mengenai bagaimana untuk menggunakan visualisasi yang telah digunakan untuk mendapatkan informasi dari sekumpulan data dan model *data mining* terkait dengan jawaban atas pertanyaan-pertanyaan bisnis yang ada.

2.11.2 *Verify and Present the Visualization or Mining Model* (Verifikasi Dan Mempresentasikan Model Visualisasi)

Pada tahap terakhir ini menjelaskan tentang tiga bagian yang ada di dalam tahap ini yaitu: memverifikasi visualisasi dan model *data mining* telah memenuhi tujuan bisnis dan objektifnya, mempresentasikan temuan dari hasil visualisasi dan *data mining* kepada pemegang keputusan, dan jika telah sesuai, dijalankan dalam lingkungan produksi.

Verifikasi visualisasi dan model *data mining* dimulai dengan mempertimbangkan transformasi *data set* yang digunakan untuk visualisasi.

Untuk mempresentasikan visualisasi harus terdiri dari tiga bagian utama yaitu deskripsi dari VDM *project goal* dan *objective*, pokok dari temuan dalam proyek untuk menyikapi atas pertanyaan bisnis yang ada, dan mendiskusikan langkah selanjutnya atau *call to action*.

