



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, 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 III

METODOLOGI PENELITIAN

3.1 Objek Penelitian

Objek penelitian yang digunakan adalah laporan GDP tahunan yang berasal dari *World Bank* dari periode 2010-2018. Data GDP yang digunakan berupa *GDP Current* yang berisikan angka riil pendapatan GDP negara dalam satuan US Dollar beserta pertumbuhannya di dalam data *GDP Growth*, *GDP inflation* untuk mengukur tingkat inflasi pada suatu negara setiap tahunnya dalam bentuk presentase, dan data *GDP per Capita* untuk mengukur pendapatan GDP per kapita setiap negara serta data *GDP per Capita Growth* untuk mengukur tingkat pertumbuhan *GDP per Capita* negara dalam bentuk presentase.

3.2 Pengumpulan Data

Pengumpulan data dilakukan dengan mengumpulkan data data terkait yang dibutuhkan terkait visualisasi GDP yang sebagian besar berasal dari data laporan tahunan GDP dari *World Bank* periode 2010-2018 yang dapat diakses secara gratis melalui bank data *World Bank* yang beralamatkan di situs data.worldbank.org

3.3 Metode Pengolahan Data

Metode pengolahan data yang akan digunakan dalam penelitian ini adalah menggunakan metode *Visual Data Mining*. Alasan penggunaan metode *Visual Data Mining* karena memiliki beberapa keunggulan ketimbang metode data mining secara umum. Menurut (Bhatt, 2016) *Visual Data Mining* memiliki beberapa

keunggulan ketimbang metode *data mining* secara umum seperti penjelesan pada tabel berikut:

<i>Visual Data Mining</i>	<i>Data Mining</i> secara umum
<i>Visual data mining</i> dapat mengeksplorasi data yang memiliki karakteristik beragam dan memiliki jumlah data yang banyak dalam bentuk sebuah visualisasi yang lebih menarik dan mudah dimengerti	<i>Data Mining</i> secara umum hanya dapat menampilkan data secara statistik dan relatif sulit untuk dimengerti
<i>Visual data mining</i> tidak membutuhkan algoritma yang kompleks ketimbang metode <i>data mining</i> lainnya	Metode <i>data mining</i> secara umum membutuhkan algoritma pemrosesan yang kompleks dan relatif sulit untuk digunakan.
<i>Visual Data Mining</i> dapat memberikan langsung berupa gambaran inti informasi yang ingin disampaikan dalam data	<i>Data Mining</i> secara umum cukup sulit untuk secara langsung menyampaikan apa inti informasi yang ingin disampaikan dalam data

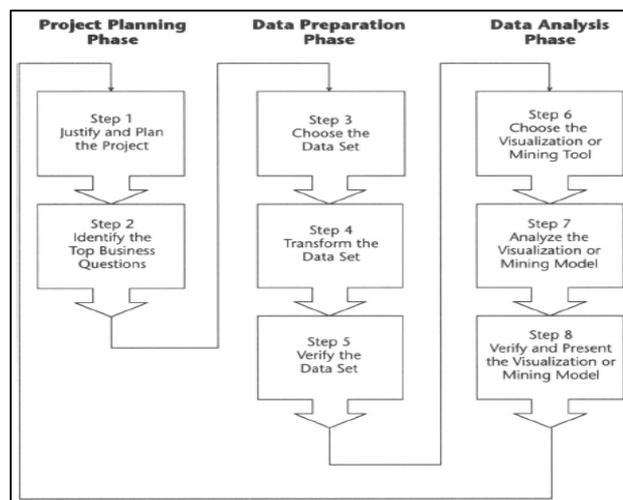
Tabel 3.1 Tabel Perbandingan metode yang digunakan

Pada Tabel 3.1 merupakan keunggulan metode pengolahan data *Visual Data Mining* ketimbang pengolahan *Data Mining* secara umum menurut (Bhatt, 2016) dimana *Visual Data Mining* memilkik banyak keunggulan seperti dapat mengeksplorasi data dengan jumlah banyak ke dalam bentuk sebuah visualisasi yang lebih mudah dimengerti, algoritma *Visual Data Mining* yang lebih sederhana dan tidak kompleks serta dapat memberikan inti informasi yang ingin disampaikan pada data yang ada. Dengan banyak keunggulan dari metode *Visual Data Mining*

ketimbang metode *Data Mining* secara umum tersebut maka *Visual Data Mining* dinilai lebih tepat untuk diterapkan ke dalam penelitian ini.

3.4 *Visual Data Mining*

Di dalam metode pengolahan data dengan metode *Visual Data Mining* secara umum terdapat delapan langkah yang terdiri dalam tiga fase atau prosedur yang harus dilakukan (Soukup & Davidson, 2002).



Gambar 3.1 Delapan tahapan *Visual Data Mining*

Sumber: (Soukup & Davidson, 2002)

Gambar 3.1 merupakan gambaran dari proses yang harus dilalui ketika melakukan melakukan metode *Visual Data Mining* menurut (Soukup & Davidson, 2002)

3.5 *Project Planning Phase*

Fase pertama dalam memulai visualisasi menggunakan metode *Visual Data Mining* adalah fase *Project Planning*. Di dalam fase ini, kita dimulai dengan menentukan bagaimana *project* yang akan digarap akan berjalan mulai dari

menentukan *tools* yang akan digunakan hingga sumber data yang akan dipakai apakah dapat digunakan untuk visualisasi ataupun tidak. Segala bentuk rencana untuk melakukan visualisasi kedepannya akan dirancang di dalam fase ini untuk memastikan proses berikutnya dapat berjalan dengan baik. Terdapat dua tahapan di dalam fase ini yaitu; *Justify and Plan the Project* dan *Identify the Top Business Questions*.

3.5.1 *Justify and Plan The Project*

Tahap pertama dalam fase *Project Planning Phase* yaitu *Justify and Plan the project*. Dalam tahap ini mengharuskan untuk menentukan tipe *project* yang akan dikerjakan. Terdapat 3 tipe *project* yang dapat dikerjakan yaitu:

- *Proof-of-concept VDM Project* memiliki *scope* yang terbatas. Secara keseluruhan *scope* di dalam *Proof-of-concept VDM Project* adalah untuk menentukan apakah hasil visualisasi yang dihasilkan memiliki keuntungan untuk bisnis, dan membantu pengambilan keputusan dari hasil visualisasi. Data *sample* yang digunakan relatif *sample* data yang kecil dan durasi *project* dilakukan hanya beberapa minggu saja
- *Pilot VDM Project* juga memiliki *scope* yang terbatas juga. *Scope Pilot VDM Project* adalah untuk menginvestigasi, menganalisa, dan untuk menjawab *bussines question* yang ada. Durasi *project* dilakukan dalam beberapa bulan.
- *Production VDM Project* memiliki *scope* yang kurang lebih mirip dengan *Pilot VDM Project* namun hasil visualisasi akan digunakan dan diimplementasikan

langsung ke dalam lingkungan produksi. Durasi *project* dilakukan dapat dikerjakan dalam beberapa tahun.

3.5.2 Identify the Top Business Questions

Tahap kedua dalam fase *Project Planning Phase* adalah *Identify the Top Business Questions*. Pada tahap ini, untuk melakukan sebuah visualisasi kita harus dapat memetakan *Business Question* dan Objek penelitian/*project* ke dalam *problem definition* yang dapat dianalisa lewat visualisasi. Tahap ini bertujuan untuk memastikan *output* yang dihasilkan dari visualisasi dan *tools* yang digunakan sesuai dengan *business questions*.

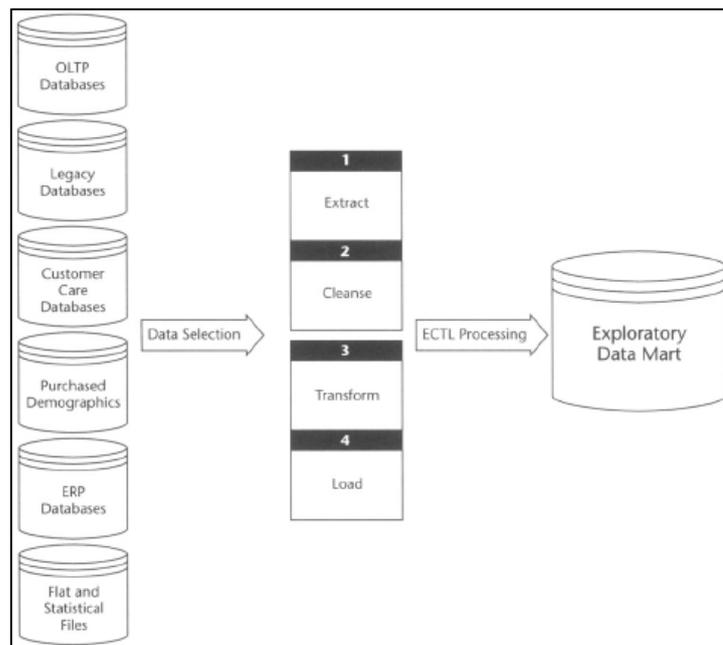
3.6 Data Preparation Phase

Fase selanjutnya adalah menyiapkan data yang akan digunakan untuk kepentingan visualisasi. Fase *Data Preparation* bertujuan untuk menemukan tipe data apa yang sesuai untuk digunakan sebagai basis visualisasi serta menyiapkan apa saja elemen yang dibutuhkan dalam data yang akan digunakan untuk visualisasi. Di dalam fase *Data Preparation* ini memiliki tiga tahapan untuk dilakukan yaitu *Choosing the Dataset*, *Transforming Data Set* dan *Verify Data Set*.

3.6.1 Choose the Data set

Untuk memberikan hasil visualisasi yang baik, tentu-nya sumber data yang akan digunakan dan dipilih untuk visualisasi haruslah sesuai dengan semua kebutuhan visualisasi yang dapat dilihat dari apa saja elemen pada data yang bisa atau tidak dapat untuk digunakan, dan juga pertimbangan lainnya seperti tipe data yang bisa digunakan dan juga *scope* ukuran data yang akan diolah nanti-nya.

Kumpulan semua data yang dibutuhkan untuk keperluan visualisasi akan dikumpulkan menjadi sebuah *Exploratory Data Mart* yang akan digunakan sebagai basis data visualiasi.



Gambar 3.2 Tahapan *Transform Data Set*

(Soukup & Davidson, 2002)

3.6.2 Transform the Dataset

Setelah data yang akan digunakan telah terkumpul ke dalam *Exploratory Data Mart*, maka kita dapat melakukan beberapa transformasi pada data tersebut agar nanti-nya dapat lebih mudah dicerna dan diolah oleh *tools* yang akan dipakai. Proses transformasi data memiliki 2 proses yang dapat digunakan untuk memastikan transformasi data berjalan dengan baik, yaitu;

a.) *Table-Level Logical Transformation*

Transformasi ini mempengaruhi struktur dan jumlah *row* di dalam sebuah data set. Berikut langkah-langkah *Table-Level Logical Transformation* yang biasa digunakan:

- *Transforming weighted data sets* merupakan transformasi yang digunakan untuk merancang dataset dengan menambahkan row baru berdasarkan kolom atau record.
- *Transforming time series data sets* merupakan transformasi yang digunakan untuk merancang dataset dimana setiap kolom di dalam dataset mewakili dimensi waktu tertentu.
- *Aggregating datasets* merupakan transformasi yang digunakan untuk menggabungkan beberapa dataset untuk mencari data statistik yang diperlukan.
- *Filtering Datasets* merupakan transformasi yang digunakan untuk menyaring elemen pada data yang bertujuan untuk memaksimalkan elemen pada data yang akan digunakan untuk visualisasi.

b.) *Column-Level Logical Transformation*

Transformasi ini mempengaruhi struktur dan jumlah kolom di dalam sebuah data set. Berikut langkah-langkah *Column-Level Logical Transformation*:

- *Simple column transformation* merupakan transformasi pada kolom secara sederhana seperti penghapusan kolom, perubahan tipe data atau penambahan kolom baru sesuai kebutuhan visualiasasi.

- *Column value grouping transformation* merupakan transformasi pada kolom yang digunakan beberapa kolom sesuai *value*-nya yang bertujuan untuk mencari pattern yang sesuai.

3.6.3 Verify the Dataset

Pada tahap ini, data yang telah dipilih dan juga melewati proses transformasi akan diverifikasi kembali apakah data yang akan digunakan sudah dapat digunakan dan juga sesuai serta akurat dan juga bebas dari *error* ketika akan digunakan untuk visualiasi. Apabila semua data yang akan digunakan telah lengkap bebas dari *error* maka data telah siap untuk digunakan ke proses selanjutnya.

3.7 Data Analysis Phase

Fase ketiga dalam metode *Visual Data Mining* adalah *Data Analysis Phase*. Fase ini bertujuan untuk menentukan *tools* apa saja yang baik digunakan untuk visualisasi berdasarkan data yang ada, dan memastikan hasil visualiasi yang didapat sesuai dengan data dan siap untuk ditampilkan. Fase ini memiliki tiga tahapan yang akan dilewati yaitu; *Choose the Visualization or Mining Tool*, *Analyze the Visualization or Mining Model* dan *Verify and Present the Visualization or Mining Model*.

3.7.1 Choose the Visualization or Mining Tool

Pada tahap ini, kita dapat memilih bentuk visualisasi apa yang cocok untuk digunakan sebagai visualisasi. Pengertian *tools* dalam visualisasi ini adalah bentuk gambaran visualisasi yang dirasa tepat dan dapat digunakan untuk dikembangkan

berdasarkan data yang didapat untuk menjawab *business question* sebelum-nya.

Untuk *tools* yang digunakan adalah Tableau

3.7.2 Analyze the Visualization

Pada tahap kedua fase *Data Analysis* adalah *Analyzing the Visualization*. Di dalam tahap ini, kita dapat menggali informasi yang dibutuhkan berdasarkan hasil visualisasi yang telah dibuat sebelumnya. Hasil analisa berupa informasi yang didapat dari gambaran visualisasi yang ada untuk menjawab *business question* yang ada.

3.7.3 Verify and Present the Visualization

Pada tahap terakhir yaitu melakukan verifikasi yang berupaya untuk memastikan hasil visualisasi yang telah dikerjakan dapat digunakan dengan baik dan seluruh tampilan yang ada di dalam visualisasi dapat diterima dengan baik.