



### **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 III

### METODOLOGI PENELITIAN

#### 3.1 Objek Penelitian

Penelitian ini dilakukan dengan menggunakan data kesehatan yang didapatkan dari portal resmi Badan Pusat Statistik ([www.bps.go.id](http://www.bps.go.id)) yaitu data indikator penyakit (Demam Berdarah, Malaria, Polio), ketersediaan rumah sakit dan tenaga medis.

Badan Pusat Statistik adalah Lembaga Pemerintahan Non Kementrian yang bertanggung jawab langsung kepada Presiden. Sebelumnya, Badan Pusat Statistik merupakan Biro Pusat Statistik, yang dibentuk berdasarkan UU Nomor 6 Tahun 1960 tentang Sensus dan UU Nomor 7 Tahun 1960 tentang Statistik. Sebagai pengganti kedua UU tersebut ditetapkan UU Nomor 16 Tahun 1997 tentang Statistik. Berdasarkan UU ini yang ditindak lanjuti dengan peraturan perundangan dibawahnya, secara formal nama Biro Pusat Statistik diganti menjadi Badan Pusat Statistik.

Berdasarkan undang-undang yang telah disebutkan di atas, peranan yang harus dijalankan adalah sebagai berikut :

- Menyediakan kebutuhan data bagi pemerintah dan masyarakat. Data ini didapatkan dari sensus atau survey yang dilakukan sendiri dan juga dari departemen atau lembaga pemerintahan lainnya sebagai data sekunder.

- Membantu kegiatan statistik di departemen, lembaga pemerintah atau intuisi lainnya, dalam membangun sistem perstatistikan nasional.
- Mengembangkan dan mempromosikan standar teknik dan metodologi statistik, dan menyediakan pelayanan pada bidang pendidikan dan pelatihan statistik.
- Membangun kerjasama dengan intuisi internasional dan negara lain untuk kepentingan perkembangan statistik Indonesia.

Gambar 3.1 menunjukkan artikel mengenai Cara baru dalam meningkatkan akses pelayanan kesehatan di seluruh kabupaten/kota Indonesia dalam *website* ([www.depkes.go.id](http://www.depkes.go.id)) .

UMMN

## 3.2 Metodologi Penelitian

Metodologi penelitian yang digunakan dalam penelitian ini terdiri dari beberapa tahap :

### 3.2.1 Studi Literatur

Tahapan ini dilakukan dengan mencari berbagai referensi dari buku-buku yang tersedia di perpustakaan, jurnal, *website* dan juga tulisan ilmiah yang berhubungan dengan penelitian yang terkait dengan visualisasi data dan *Tableau*.

### 3.2.2 Pengumpulan Data

Tahapan pengumpulan data dilakukan dengan mencari data pada bidang kesehatan yang dapat divisualisasikan dalam penelitian ini. Data yang digunakan berasal dari *website* ([www.bps.go.id](http://www.bps.go.id)). Data yang digunakan yakni data tahun 2012-2015 karena pada data tahun 2016 masih belum lengkap sehingga tidak memungkinkan untuk divisualisaikan selain itu data kesehatan yang di gunakan hanya dapat di dapatkan dalam jangka waktu tertentu. Jumlah data keseluruhan yang digunakan untuk melakukan visualisasi data pada file *Ms Excel* dan *csv* berjumlah 10.200 *field*. Tabel 3.1 menunjukkan data yang digunakan untuk visualisasi data indikator penyakit, ketersediaan rumah sakit dan tenaga medis.

Tabel 3.1 Data Kesehatan (data per tahun 2012-2015)

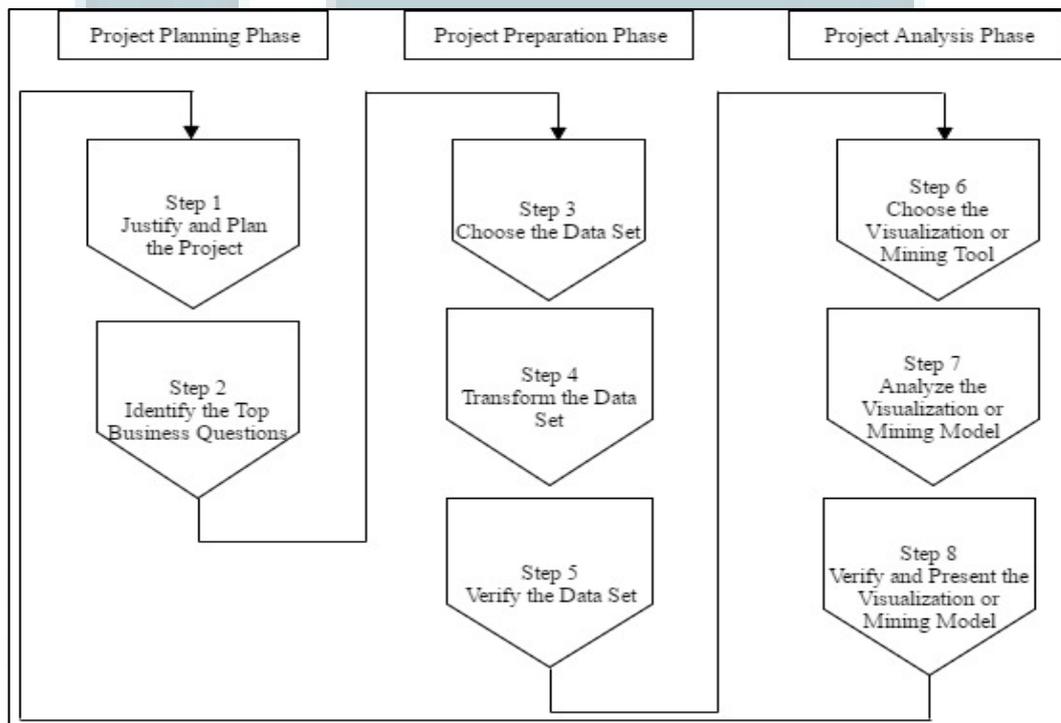
Sumber Data	Nama Variabel	Keterangan Variabel
Data kesehatan	kode_provinsi	Kode_Provinsi
	nama_provinsi	Nama Provinsi
	nama_kabupaten	Nama Kabupaten
	indikator_Penyakit	Indikator Penyakit
	nilai	Nilai yang terkait dengan indikator penyakit
	jenis_rumah_sakit	Jenis dari Rumah Sakit (RS Umum)
	jumlah_rumah_sakit	Jumlah Rumah Sakit
	jumlah_tempat_tidur	Jumlah Tempat Tidur
	jenis_tenaga	Jenis Tenaga SDM Kesehatan
	jumlah_tenaga	Jumlah Tenaga Terkait
	lintang	Garis lintang
	bujur	Garis bujur

### 3.2.3 Konsultasi

Dalam metode ini penulis juga melakukan konsultasi data ke Badan Pusat Statistik dengan tujuan agar data yang sedang diteliti yaitu data indikator penyakit, ketersediaan rumah sakit dan tenaga medis, apakah tersedia di Badan Pusat Statistik dan apakah data tersebut dapat digunakan dan sesuai dengan kriteria data yang dibutuhkan dalam memenuhi penelitian yang sedang penulis kerjakan.

### 3.3 Metode Visual Data Mining (VDM)

Metode penelitian yang digunakan adalah *visual data mining* (VDM). *Visual data mining* merupakan sebuah metode yang digunakan untuk menemukan pola data yang belum diketahui sebelumnya, pola tingkah laku data yang ada di dalam perusahaan atau instansi terkait nantinya akan sangat membantu dalam melakukan penelusuran pada data yang digunakan melalui teknik visual, sehingga instansi atau perusahaan terkait dapat terbantu dan memperoleh pengetahuan serta pemahaman yang mendalam mengenai data-data yang tersedia dalam bentuk visual selain itu juga membantu perusahaan dalam mengambil keputusan-keputusan yang sulit.



Gambar 3.1 Eight-step data Visualization and Visual data mining methodology

Sumber : *Visual Data Mining : techniques and Tools for Data Visualization and Mining*

Gambar 3.2 memperlihatkan tahapan-tahapan bagaimana metode VDM (*Visual Data Mining*) dalam membangun visualisasi ini. Detail tahapan akan dijelaskan lebih lanjut dalam sub-sub bab di bawah ini.

### 3.4 *Project Planning Phase*

Tahap ini merupakan langkah paling pertama dalam melakukan analisa terhadap perencanaan dan pembangunan desain visualisasi data, dalam tahap ini juga menjelaskan bagaimana visualisasi dari data *mining model* membantu dalam menganalisa bisnis serta membantu dalam melakukan *decision making*. Tidak hanya untuk menganalisa dalam tahap ini juga akan menjelaskan bagaimana penggunaan *visualization tools* dapat memaksimalkan potensi dalam penyebaran dari model tersebut.

#### 3.4.1 *Justify and Plan the Project*

Tahap *justify and plan the project* merupakan langkah-langkah dalam merencanakan proyek dalam membuat visualisasi. Ada 3 tipe dari proyek data visualisasi yaitu:

- *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 bisnis dalam perusahaan untuk membuktikan kepada pembuat keputusan.
- *A pilot VDM project* memiliki *scope* yang terbatas, secara keseluruhan *scope*-nya untuk menginvestigasi, menganalisis, dan menjawab satu atau banyak *business question*.

- A *production VDM project* scope-nya memiliki kemiripan dengan *pilot project*, namun dalam hal visualisasi dan model *data mining* yang hasilnya diimplementasikan ke dalam lingkungan produksi. Jadi secara keseluruhan scope-nya adalah untuk melakukan investigasi, menganalisis dan menjawab satu atau banyak *business question* dan mengimplementasikannya dan mengukur hasil dari model visualisasi produksi dan *data mining* yang telah dibuat.

#### 3.4.2 *Identify the Top Business Questions*

Dalam langkah ini kita harus memastikan *business question* apa yang dibutuhkan untuk membangun model data visualisasi dan *data mining*, seperti memetakan *business question* ke dalam *problem definition* yang dapat diatasi dengan model data visualisasi dan *data mining*. Selain itu harus menentukan *output* dan memperjelas apa yang diharapkan dari proyek ini.

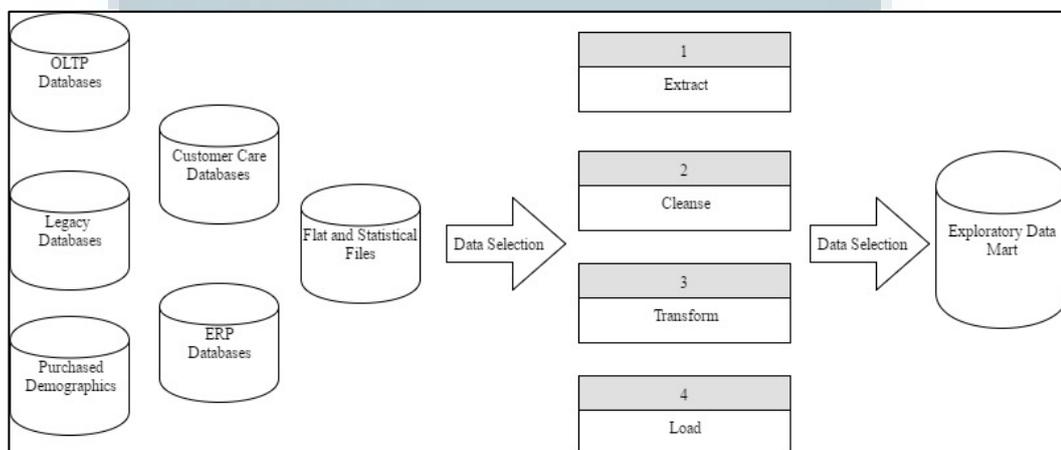
### 3.5 *Data Preparation Phase*

Tahap kedua ini merupakan tahapan dimana untuk membangun proyek visualisasi yang akan dibangun, hal-hal apa saja yang diperhatikan seperti dalam hal pemilihan data yang sesuai agar tidak keluar dari tujuan utama pembangunan proyek visualisasi dan tidak lupa data yang akan digunakan merupakan data yang sudah valid.

#### 3.5.1 *Choose the Data Set*

Untuk menjawab tahap *top business questions* tentu harus memiliki *data set* yang valid dan dapat digunakan, mendapatkan data-data yang akan digunakan

bukanlah hal yang mudah untuk dilakukan karena dalam membangun data visualisasi hal terpenting merupakan data-data yang menjadi inti dari tahap ini, sehingga hal pertama yang dilakukan dalam membangun sebuah data visualisasi adalah mengetahui dimana sumber data bisa didapatkan, lalu bagaimana cara mendapatkannya, kemudian jenis atau tipe data seperti apa yang dibutuhkan untuk memenuhi kebutuhan dalam visualisasi data dan yang terakhir berapa banyak data yang akan digunakan. Setelah data-data telah berhasil terkumpul maka langkah selanjutnya adalah melakukan desain *exploratory data mart* dengan menggunakan ECTL untuk mengisi *exploratory data mart* yang ada tujuannya untuk menyatukan informasi dari sumber data yang berbeda-beda ke dalam satu *exploratory data mart* dan dapat digunakan dalam mengambil keputusan.



Gambar 3.2 *Data flow from operational data sources to the visualization and data mining tools*

Sumber : *Visual Data Mining: techniques and Tools for Data Visualization and Mining* (Soukup, 2002)

### 3.5.2 *Transform the Data Set*

Langkah berikut adalah proses peningkatan informasi dari data yang digunakan, penambahan informasi ke dalam data bisa dilakukan dengan banyak cara bahkan sampai mengganti format data tersebut untuk mendukung pembangunan data visualisasi, proses mining dan juga proses investigasi data. Dalam tahap ini juga dilakukan *cleansing* yang ada di dalam data, *cleansing* ini bertujuan agar data yang akan digunakan dalam visualisasi akan lebih akurat dan terhindar dari kesalahan penyampaian informasi setelah data tersebut sudah divisualisasi. Ada 2 proses yang akan dilakukan sebelum melakukan visualisasi data, yakni :

a) *Table Level Logical Transformation* :

- *Transforming weighted data sets* merupakan cara membuat *data set* baru dengan menambahkan kolom baru berdasarkan pembobotan kolom atau *record*
- *Transformation time series* merupakan cara membuat *data set* dimana setiap kolom mewakili dimensi waktu tertentu. Dalam proses beberapa data visualisasi akan menggunakan *time series*.
- *Aggregating data sets* merupakan cara menggabungkan sekumpulan data, seperti menggabungkan data sebaran penyakit dengan ketersediaan rumah sakit dan tenaga medis sehingga dapat membantu menjawab *business questions*

- *Filtering data sets*: melakukan penyaringan *data set* berdasarkan *filtering condition* dengan tujuan agar data yang di visualkan akan menjadi lebih akurat

b) *Column Level Logical Transformation*:

- *Simple column transformation* merupakan transformasi menghapus, merubah atau membuat kolom baru. Transformasi ini memungkinkan untuk memanipulasi *data set* dengan cara menggabungkan banyak kolom ke dalam satu kolom yang akan digunakan dalam visualisasi sehingga meningkatkan akurasi dari model *data mining*
- *Grouping column* merupakan gabungan kolom yang sudah ada ke dalam *grouping* yang lebih besar
- Membuat agregasi kolom untuk meningkatkan akurasi *data mining model*

### 3.5.3 *Verify the Data Set*

Dalam tahap ini dilakukan identifikasi dan pengujian terhadap *data set* yang telah melewati tahap ECTL untuk memastikan tidak terdapat data yang error atau bias. Untuk meminimalisir terjadinya tidak akuratnya data yang akan di visualisasikan perlu adanya proses pengulangan ECTL hal ini diharapkan agar data yang akan digunakan dapat meningkatkan akurasi data yang ada. Proses pengulangan ECTL merupakan hal wajar dalam membangun sebuah data visualisasi.

Data visualisasi merupakan hal yang sangat berhubungan erat dengan data yang ada maka kemungkinan data yang kurang atau data tersebut merupakan data

lama dalam rentang waktu yang lama, jadi merupakan hal yang biasa bila pada saat pertengahan pengerjaan dalam membuat visualisasi data akan terjadi perubahan dalam *data set*.

### **3.6 *Project Analysis Phase***

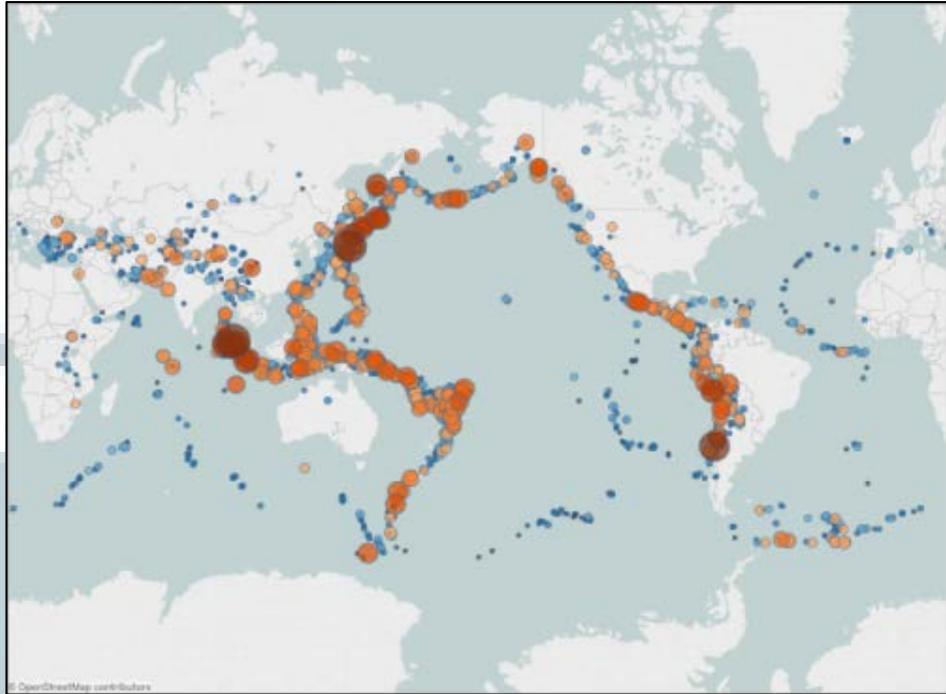
Pada tahap ketiga dalam metode VDM ini akan dilakukan analisa bagaimana cara memilih *tools* yang tepat dan dapat digunakan berdasarkan analisa yang telah dilakukan, langkah selanjutnya adalah melakukan pengecekan akhir terhadap visualisasi yang telah dibangun agar bebas dari *error* dan juga bias, serta langkah terakhir yang dilakukan adalah mempersiapkan tahap persentasi model visualisasi yang sudah dibangun.

#### **3.6.1 *Choose the Visualization or Mining Tools***

Dalam tahap ini untuk membangun sebuah data visualisasi maka dibutuhkan sebuah *tools* untuk mendukung pengerjaan dalam proses memvisualkan data-data yang ada dan sesuai dengan *business questions* yang telah diteliti. *Data Visualization tool* dapat dibagi menjadi 3 kategori yaitu :

U  
M  
N

a. *Proportional Symbol Maps*



Gambar 3.3 *Proportional Symbol Maps*

Sumber : ([www.tableau.com](http://www.tableau.com))

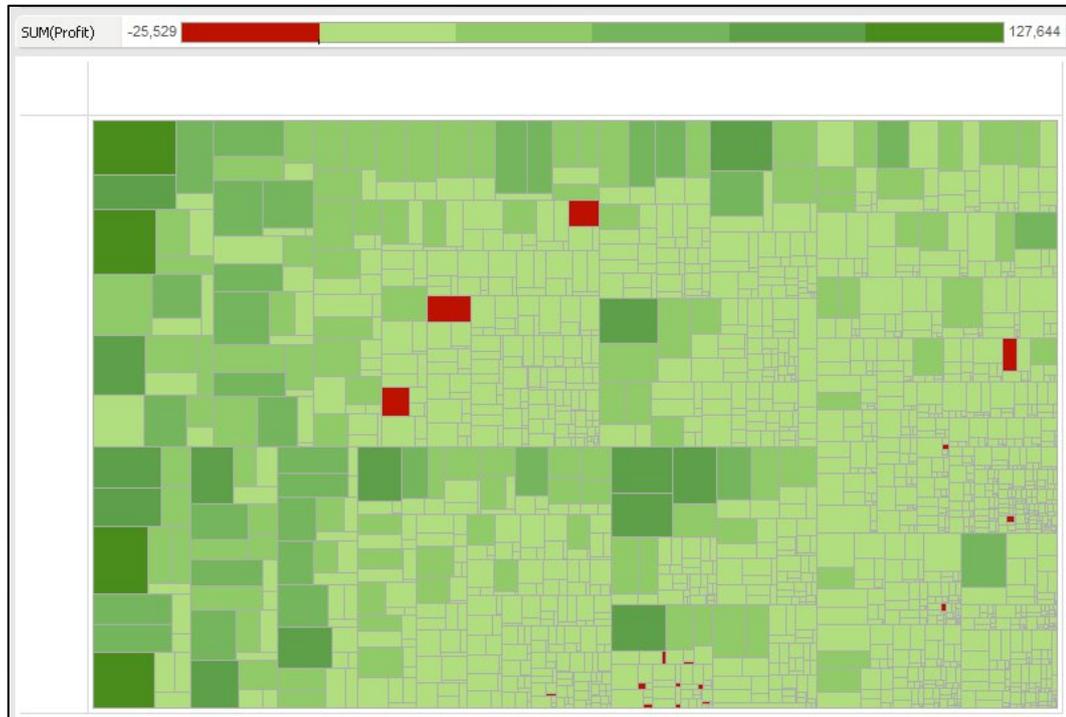
*Proportional symbol maps* dapat digunakan untuk melakukan analisa berdasarkan data kuantitatif untuk setiap lokasi.

### 3.6.2 *Analyze the Visualization or Mining Model*

Dalam tahap ini juga dilakukan analisa terhadap model visualisasi yang dipilih apakah akan lebih berguna dibandingkan dengan model visualisasi yang lainnya dalam menjawab pertanyaan bisnis.

Berikut ini adalah beberapa data visualisasi yang dapat digunakan dalam menganalisa dan mengevaluasi *data set* dan mendapatkan informasi dari pertanyaan bisnis yang ada :

a. *Treemap Chart*

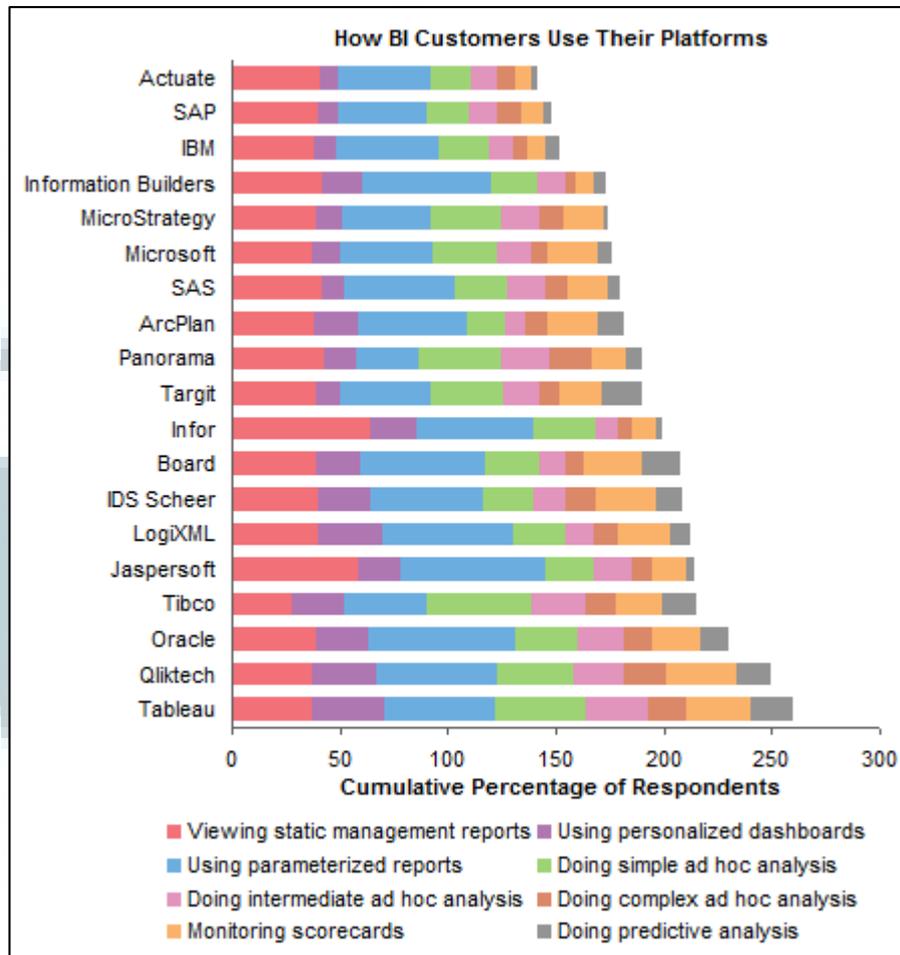


Gambar 3.4 *Treemap Chart*

Sumber : (<https://www.interworks.com/blog/ccapitula/2014/10/14/tableau-essentials-chart-types-treemap>)

Treemap Chart adalah metode grafis yang menampilkan *data multivariant* dalam bentuk hirarki dua dimensi sangat berguna untuk menampilkan data yang berjumlah ratusan maupun ribuan.

b. *Stacked Bar Chart*

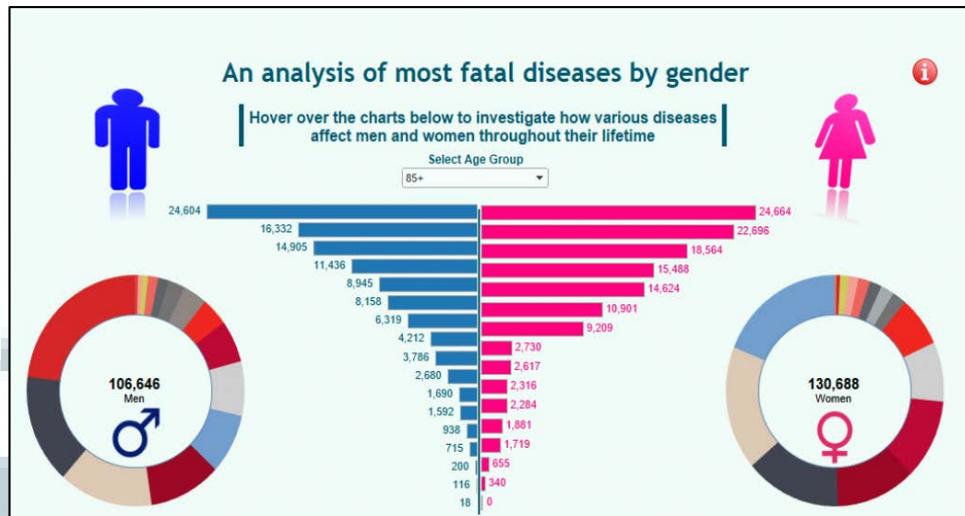


Gambar 3.5 *Stacked Bar Chart*

Sumber : (<http://peltiertech.com/stacked-bar-chart-alternatives/>)

*Stacked bar chart* dapat digunakan untuk melihat pola trend atau perkembangan setiap tahun dengan cara menambahkan lapisan lain ke dalam *bar chart*.

c. *Donut Chart*



Gambar 3.6 *Donut Chart*

Sumber : (www.concentra.co.uk)

*Donut Chart* berfungsi untuk mempresentasikan data dalam bentuk persentase.

### 3.6.3 *Verify and Present the Visualization or Mining Model*

Merupakan tahap terakhir metode VDM dalam *data analysis phase*, dalam langkah ini merupakan gabungan dari tiga bagian yaitu :

1. Melakukan verifikasi bahwa model visualisasi yang dibuat sudah sesuai dengan *top business questions* dan memenuhi tujuannya.
2. Melakukan persiapan persentasi dari model visualisasi data yang telah dibuat atas temuan yang ada di dalamnya untuk diberikan kepada *user* atau pengambil keputusan.
3. Melakukan publish visualisasi data.

### 3.7 Penelitian Terdahulu

Penelitian terdahulu merupakan penelitian yang telah dilakukan sebelumnya dalam ruang lingkup topik pembahasan yang serupa dengan topik yang akan diteliti saat ini. Penelitian terdahulu sangat diperlukan dengan tujuan agar peneliti dapat memahami dan mengerti perbedaan mengenai tujuan dan fokus pada penelitian yang dilakukan.

Oleh karena itu, sebagai bahan acuan dan pembanding, maka peneliti melakukan perbandingan terhadap tiga peneliti terdahulu dimana peneliti pertama dan kedua meneliti objek yang sama sedangkan untuk peneliti ketiga menggunakan *tools* yang sama

Penelitian pertama adalah penelitian yang dilakukan oleh **Nielza Atthina, Lizda Iswari** Universitas Islam Indonesia, Yogyakarta tahun 2014 dengan judul **“Klasterisasi Data Kesehatan Penduduk Untuk Menentukan Rentang Derajat Kesehatan Daerah dengan Metode K-Means”**. Tujuan dari penelitian ini adalah mengklaster atau mengelompokkan kabupaten-kabupaten di Provinsi Jawa Tengah dan Yogyakarta berdasarkan kemiripan nilai lima buah ukuran indikator mortalitas derajat kesehatan daerah, yaitu angka kelahiran kasar, angka kematian balita dan angka kematian ibu.

Metode yang digunakan oleh penelitian sebelumnya adalah K-Means. K-Means merupakan algoritma yang bekerja dengan cara membagi data dalam sejumlah klaster untuk dianalisis faktor kesamaan (*similarity*) maupun ketidaksamaan (*dissimilarity*) yang melekat pada pada kumpulan data tersebut dan kemudian dianalisis pola keterhubungan antar data.

Hasil dari penelitian ini berupa tampilan antar muka yang menampilkan peta klustering berdasarkan data kesehatan pada Provinsi Jawa Tengah dan Yogyakarta disertai dengan filter kordinat untuk menentukan posisi dari suatu daerah yang akan dipilih nantinya, yang nantinya akan menampilkan hasil dari klusterisasi data indikator kesehatan dalam bentuk warna hijau, merah dan kuning yang mewakili jumlah angka kematian atau kehidupan pada Provinsi Jawa Tengah.

Penelitian kedua adalah yang dilakukan oleh **Surya Afnarius** Teknik Elektro Universitas Andalas, Padang tahun 2008 dengan Judul “**Perancangan Sistem Visualisasi Data Kepengungsian Menggunakan Database Spatial POSTGIS**”. Tujuan dari penelitian ini adalah untuk membantu para penyumbang atau pendonor untuk mengetahui lokasi pengungsian dan apa yang dibutuhkan oleh para pengungsi dengan menggunakan database spatial PostGIS.

Metode yang digunakan dalam penelitian ini adalah metode *Waterfall*, karena mempertimbangkan kelebihan dan kelemahan jenis pekerjaan pada sistem GIS. Pengembangan sistem GIS cenderung dilakukan dengan menggunakan model yang bersifat dinamis dan dapat mengakomodasikan perubahan-perubahan kebutuhan secara fleksibel dengan memperhatikan faktor-faktor resiko yang mungkin dihadapi.

Hasil penelitian terdahulu berupa tampilan antar muka lokasi pengungsi berdasarkan id atau nama tempat pengungsian, lokasi pengungsian berdasarkan jumlah pengungsi, lokasi pengungsian berdasarkan kebutuhan dan lokasi fasilitas pendukung pengungsian pada area bencana x.

Penelitian ketiga adalah penelitian terdahulu yang menggunakan *tools* yang sama dengan penelitian ini. Penelitian dilakukan oleh **Cavin** Fakultas Teknik & Informatika Universitas Multimedia Nusantara, Tangerang tahun 2016 dengan Judul **“Data Visualisasi lokasi dan ratio tempat tidur rumah sakit dengan sasaran sensus penduduk di daerah Jakarta”**. Tujuan dari penelitian ini adalah membangun sebuah data visualisasi sebagai pendukung dalam proses pengambilan keputusan untuk menentukan penambahan rumah sakit baru, daerah mana yang akan menjadi prospek rumah sakit baru, menampilkan alamat rumah sakit, menampilkan jumlah tempat tidur yang dimiliki setiap rumah sakit dan perbandingan jumlah tempat tidur yang dimiliki dengan jumlah sasaran sensus penduduk tahun 2016.

Metode yang digunakan oleh peneliti terdahulu adalah metode *Visual Data Mining*. *Visual Data Mining* adalah sebuah teknik yang dapat digunakan untuk menemukan kecenderungan dalam data yang belum diketahui sebelumnya, tingkah laku dan juga anomali di dalam data perusahaan atau instansi yang nantinya akan sangat membantu perusahaan dalam melakukan penelusuran pada data memperoleh pengetahuan dan pengertian yang mendalam atas data yang ada.

Hasil penelitian terdahulu berupa visualisasi data yang dapat menampilkan *mapping* lokasi rumah sakit, puskesmas lalu total tempat tidur rumah sakit (*highlight bar chart*) kemudian detail tempat tidur rumah sakit (*side by side bar chart*) dan yang terakhir adalah rasio tempat tidur rumah sakit dan sasaran sensus tahun 2016 (*side by side bar chart*).

### 3.8 Perbandingan Tools

Perbandingan *tools* digunakan untuk mengetahui keunggulan dan kekurangan terhadap *tools* yang akan digunakan dalam melakukan visualisasi data, digunakan nya perbandingan untuk mendapatkan hasil yang lebih akurat. *Tools* yang akan dibandingkan adalah *Tableau*, Power BI, Pentaho dan QlikView. Perbandingan dapat dilihat pada Tabel 3.2 di bawah ini

Tabel 3.2 Perbandingan Tools

Sumber : (<http://reviews.financesonline.com>)

<b>Tools/Software</b>	<b>Tableau</b>	<b>Pentaho</b>	<b>Power BI</b>	<b>QlikView</b>
<b>Device Supported</b>				
<i>Website based</i>	✓	✓	✓	✓
<i>iPhone app</i>	✓	✓	✓	✓
<i>Android app</i>	✓	✓	✓	
<i>Windows</i>			✓	
<i>Phone app</i>				
<b>Typical Customer</b>				
<i>Freelancers</i>	✓			
<i>Small Businesses</i>	✓			✓
<b>Features</b>				
<i>API</i>	✓		✓	✓
<i>Business Intelligence</i>	✓	✓	✓	✓
<i>Data Visualization</i>	✓		✓	✓
<i>Dashboard Creation</i>	✓	✓		✓
<i>Customize Reporting</i>	✓	✓		✓

Dari hasil perbandingan yang di dapat pada Tabel 3.2 penulis dapat memilih *Tableau Software* karena *Tableau* memiliki keunggulan dalam hal *customize dashboard* interaktif sehingga memudahkan dalam hal analisis data yang banyak. Selain itu *Tableau* sendiri memiliki keunggulan dalam hal *platform*, dalam hal ini sudah di dukung *website based* sehingga memudahkan pengguna untuk mengupload hasil *dashboard* atau *story* ke dalam *website* agar dapat di *share* ke pengguna lainnya sehingga visualisasi dapat dikembangkan terus-menerus seiring berjalannya waktu.

UMMN