

**PENGGUNAAN JARAK *DYNAMIC TIME WARPING* (DTW)  
PADA ANALISIS *CLUSTER* DATA DERET WAKTU**

**(StudiKasus Pada Dana Pihak Ketiga Provinsi Se-Indonesia)**

**SKRIPSI**

Sebagai salah satu syarat untuk memperoleh gelar  
Sarjana Sains dalam bidang statistika

oleh :

**ARDIANSYAH MUHAMMAD**

**105090500111028 - 95**



**PROGRAM STUDI STATISTIKA  
JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS BRAWIJAYA  
MALANG  
2014**

**LEMBAR PENGESAHAN SKRIPSI**  
**PENGUNAAN JARAK *DYNAMIC TIME WARPING* (DTW)**  
**PADA ANALISIS *CLUSTER* DATA DERET WAKTU**  
**(Studi Kasus Pada Dana Pihak Ketiga Provinsi Se-Indonesia)**

oleh :  
**ARDIANSYAH MUHAMMAD**  
**105090500111028 - 95**

Setelah dipertahankan di depan Majelis Penguji  
pada tanggal 15 Juli 2014  
dan dinyatakan memenuhi syarat untuk memperoleh gelar  
Sarjana Sains dalam bidang statistika

Mengetahui,

Dosen Pembimbing,

**Samingun Handoyo, S.Si., M.Cs.**  
**NIP. 197304151998021002**

**Ketua Jurusan Matematika**  
**Fakultas MIPA Universitas Brawijaya**

**Dr. Abdul Rouf Alghofari, M.Sc**  
**NIP. 19670971992031001**

## LEMBAR PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : Ardiansyah Muhammad  
NIM : 105090500111028  
Jurusan : Statistika  
Penulis Skripsi berjudul :

**PENGUNAAN JARAK *DYNAMIC TIME WARPING* PADA  
ANALISIS *CLUSTER* DATA DERET WAKTU (Studi Kasus  
Pada Dana Pihak Ketiga Provinsi Se-Indonesia )**

Dengan ini menyatakan bahwa :

1. Isi dari Skripsi ini saya buat adalah benar-benar karya sendiri tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam Skripsi ini.
2. Apabila dikemudian hari ternyata Skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala yang akan saya terima.

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 15 Juli 2014  
Yang menyatakan,

**(ARDIANSYAH MUHAMMAD)**  
**NIM. 105090500111028**

# PENGUNAAN JARAK *DYNAMIC TIME WARPING* (DTW) PADA ANALISIS *CLUSTER* DATA DERET WAKTU

(Studi Kasus Pada Dana Pihak Ketiga Provinsi Se-Indonesia)

## ABSTRAK

Analisis *cluster* merupakan analisis multivariat yang bertujuan untuk mengelompokkan objek berdasarkan karakteristik yang dimiliki. Analisis *cluster* data deret waktu merupakan salah satu teknik *data mining* yang dikembangkan untuk mengelompokkan objek berdasarkan data deret waktunya. Data deret waktu digunakan di berbagai bidang diantaranya bisnis, teknik, dan ekonomi. Pengelompokan data deret waktu dilakukan dengan menggunakan jarak *Autocorrelation-based*, *Complexity Invariant*, *Periodogram-based*, dan *Dynamic Time Warping* (DTW). *Dynamic Time Warping* merupakan teknik pemrograman dinamis yang dapat menemukan jarak optimal antara dua deret waktu. Pemilihan jarak terbaik menggunakan korelasi *cophenetic* hasil analisis *cluster* hirarki menggunakan *average linkage* pada jarak *Autocorrelation-based*, *Complexity Invariant*, *Periodogram-based*, dan *Dynamic Time Warping*. Optimalisasi hasil analisis *cluster* hirarki menggunakan algoritma *K-means* dengan jumlah kelompok yang didapatkan melalui *cluster* hirarki. Validasi pengelompokan dengan algoritma *K-means* menggunakan koefisien *silhouette*. Nilai korelasi *cophenetic* menunjukkan bahwa *Dynamic Time Warping* merupakan jarak terbaik dengan nilai korelasi *cophenetic* sebesar 0,9775802. Validasi pengelompokan menggunakan algoritma *K-means* menunjukkan jarak *Dynamic Time Warping* menghasilkan pengelompokan yang cukup baik dengan nilai koefisien *silhouette* sebesar 0,257.

*Kata Kunci: Dynamic Time Warping, Time Series Clustering, Cluster Hirarki, Algoritma K-means*

# APPLICATION DYNAMIC TIME WARPING (DTW) DISTANCE ON ANALYSIS CLUSTER OF TIME SERIES DATA

(StudyOnThird Party Funds In Indonesia)

## ABSTRACT

Cluster analysis is a multivariate analysis which aims to classify objects based on their characteristics. Cluster analysis of time series data is one of the data mining techniques developed for classifying objects based on their time series structure. Time series data is used in a variety of fields including business, engineering, and economics. Grouping time series data is done by using Autocorrelation-based, Complexity Invariant, Periodogram-based, and Dynamic Time Warping (DTW) Distance. Dynamic Time Warping is a dynamic programming technique that can find the optimal distance between two time series. The best distance is selected by using cophenetic correlation results of hierarchical cluster analysis using the average linkage using Autocorrelation-based, Complexity Invariant, Periodogram-based, and Dynamic Time Warping Distance. Optimization of the results of a hierarchical cluster analysis using K-means algorithm with the number of groups obtained through the cluster hierarchy. Validation of clustering with K-means clustering algorithm use silhouette coefficient. Cophenetic correlation value generated from hierarchical cluster analysis of the Dynamic Time Warping distance is the best distance among three other distances with cophenetic correlation value of 0.9775802. Validation of clustering using K-means algorithm using Dynamic Time Warping distance show that DTW generate pretty good grouping with the silhouette coefficient value of 0.257

Keywords: Dynamic Time Warping, Time Series Clustering, Hierarchical Cluster , K-means Algorithm

## KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas berkat, rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan tugas akhir Skripsi yang berjudul “Penggunaan Jarak *Dynamic Time Warping* (DTW) Pada Analisis *Cluster* Data Deret Waktu”. Secara keseluruhan, isi dari Skripsi ini merupakan hasil penerapan ilmu Statistika. Salah satu bidang ilmu statistika yang digunakan adalah *Time Series Clustering*.

Selama pelaksanaan dan penyusunan Tugas Akhir ini telah dibantu berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada :

1. Samingun Handoyo, S.Si., M.Cs selaku Dosen Pembimbing Skripsi atas waktu dan bimbingan yang telah diberikan.
2. Dr. Suci Astutik, M.Si dan Dr. Ir. Atiek Iriany, MS selaku Dosen Penguji I dan II Skripsi atas waktu dan bimbingan yang telah diberikan.
3. Dr. Abdul Rouf Alghofari, M.Sc selaku Ketua Jurusan Matematika Program FMIPA Universitas Brawijaya.
4. Ayah dan Ibu tercinta yang telah menjadi penjaga amanah Allah terbaik, dan tak tergantikan yang pernah saya punya.
5. Adik dan keluarga tercinta yang telah merawat dan ikhlas menerima dan menemani sejak saya dilahirkan.
6. Saudari Rachmawati Dwi Agustin yang menjadi inspirasi untuk segera menyelesaikan tugas akhir ini.
7. Sahabat-sahabat Kendalsari : Edo, Hendrik, Neo, Anam, Rojo, Zainul, Rezki. Sahabat-sahabat Puncak Maleno : Faruq, Jeffri, Grandis, Ronal, Afif, Faiz, Fufah. Sahabat-sahabat Majelis Ilmu : Bima, Ashrul, Dawat, Opan yang telah ikhlas menemani berdiskusi, bermain, selama masa perkuliahan dan mendoakan kelancaran penyusunan tugas akhir ini.
8. Sahabat-sahabat Griyasantha B 106 : Miko, Danang, Raka, Vicky, Alifian, Linggo yang telah menemani selama 2 tahun masa perkuliahan
9. Teman-teman Statistika 2009, 2010, dan 2011 yang telah membantu memberikan saran serta doa dalam penyusunan tugas akhir.

Dalam penyusunan tugas akhir ini masih terdapat banyak kekurangan. Oleh karena itu kritik dan saran yang membangun akan sangat berguna demi penyusunan yang lebih baik. Semoga tugas akhir ini dapat memberikan manfaat bagi para pembaca

Malang, Juli 2014  
Penulis

UNIVERSITAS BRAWIJAYA



# DAFTAR ISI

Halaman

<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	ii
<b>LEMBAR PERNYATAAN</b> .....	iii
<b>ABSTRAK</b> .....	iv
<b>KATA PENGANTAR</b> .....	vi
<b>DAFTAR ISI</b> .....	viii
<b>DAFTAR GAMBAR</b> .....	x
<b>DAFTAR TABEL</b> .....	xi
<b>DAFTAR LAMPIRAN</b> .....	xii
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah .....	4
1.3 Tujuan .....	5
1.4 Manfaat Penelitian .....	5
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Analisis <i>Cluster</i> .....	7
2.1.1 Analisis Cluster pada Data Statis .....	6
2.1.2 Analisis Cluster pada Data Deret Waktu .....	7
2.2 Ukuran Kemiripan .....	7
2.2.1 Jarak Euclidean .....	15
2.2.2 Periodogram-based Distance .....	16
2.2.3 Complexity Invariant Distance .....	18
2.2.4 Correlation-based Distance .....	19
2.2.5 Short Time Series Distance .....	20
2.2.6 Autocorrelation-based Distance .....	21
2.3 Dynamic Time Warping .....	22
2.4 Metode Hirarki .....	25
2.5 Algoritma K-means .....	27
2.6 Pemilihan Metode Terbaik .....	27
2.6.1 Korelasi Cophenetic .....	28
2.6.2 Koefisien Silhouette .....	29
2.7 Tinjauan Non Statistika .....	34

2.7.1 Pengertian, Fungsi, dan Tujuan Bank.....	31
2.7.2 Jenis Bank.....	32
2.7.3 Sumber Dana Bank.....	33
<b>BAB III METODOLOGI</b>	
3.1 Sumber Data .....	35
3.2 Langkah-Langkah Penelitian.....	35
<b>BAB IV HASIL DAN PEMBAHASAN</b>	
4.1 Deskripsi Data .....	39
4.2 Plot Deret Waktu .....	39
4.3 Pengujian Stasioneritas Data DPK 30 Provinsi di Indonesia .....	41
4.3.1 Pengujian Stasioneritas Terhadap Ragam.....	42
4.3.2 Pengujian Stasioneritas Terhadap Rata-Rata .....	42
4.4 Stasioneritas Deret Waktu .....	42
4.5 Analisis Cluster Deret Waktu .....	44
4.5.1 Autocorrelation-based Distance.....	44
4.5.2 Complexity Invariant Distance.....	45
4.5.3 Periodogram-based Distance.....	46
4.5.4 Dynamic Time Warping.....	47
4.6 Ukuran Ketepatan Analisis Cluster Hirarki.....	48
4.7 Analisis Cluster Menggunakan Algoritma K-means.....	49
4.8 Validasi Hasil Analisis Cluster dengan Algoritma K-means.....	50
4.9 Akurasi Pengelompokan Menggunakan Jarak Dynamic Time Warping.....	51
<b>BAB V KESIMPULAN DAN SARAN</b>	
5.1 Kesimpulan .....	52
5.2 Saran .....	52
<b>DAFTAR PUSTAKA</b> .....	55
<b>LAMPIRAN</b> .....	59

## DAFTAR GAMBAR

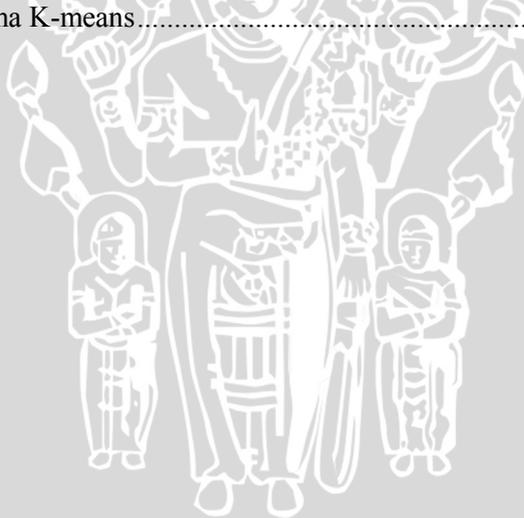
Halaman

Gambar 2.1 Pola data deret waktu dan kompleksitas .....	19
Gambar 3.1 Diagram alir analisis cluster dengan jarak periodogram-based, autocorrelation-based, CID dan DTW .....	37
Gambar 3.2 Diagram alir perhitungan jarak dynamic time warping	38
Gambar 4.1 Plot Data DPK 30 Provinsi di Indonesia .....	41
Gambar 4.2 Plot Data Dana Pihak Ketiga setelah ditransformasi ....	43
Gambar 4.3 Dendrogram analisis cluster hirarki menggunakan Autocorrelation-based Distance .....	44
Gambar 4.4 Dendrogram analisis cluster hirarki menggunakan Complexity Invariant Distance .....	45
Gambar 4.5 Dendrogram analisis cluster hirarki menggunakan Periodogram-based Distance.....	46
Gambar 4.6 Dendrogram analisis cluster hirarki menggunakan Jarak Dynamic Time Warping.....	47

## DAFTAR TABEL

Halaman

Tabel 2.1	Contoh Jarak Penggabungan Tiap Stage .....	29
Tabel 2.2	Kriteria Subjektif Kualitas Pengelompokkan berdasarkan Koefisien Silhouette .....	31
Tabel 2.3	Komposisi Dana Pihak Ketiga (DPK) Perbankan Nasional Periode 2004-2013 (Bulan Desember) dalam Milyar Rupiah .....	34
Tabel 4.1	Statistik Deskriptif DPK 30 Provinsi di Indonesia Bulan Januari 2005 - Desember 2013 dalam Milyar Rupiah.....	39
Tabel 4.2	Koefisien Korelasi Copenetic Jarak CID, DTW, Autocorrelation-based, dan Periodogram-based .....	48
Tabel 4.3	Anggota tiap cluster hasil analisis cluster menggunakan Algoritma K-means .....	50



## DAFTAR LAMPIRAN

Halaman

Lampiran 1. Data Dana Pihak Ketiga Januari 2005 – Desember 2013 30 Provinsi di Indonesia .....	59
Lampiran 2. Plot Data Dana Pihak Ketiga 30 Provinsi.....	77
Lampiran 3. Nilai Lambda Data Dana Pihak Ketiga (DPK) 30 Provinsi .....	85
Lampiran 4. Hasil Uji Augmented Dickey Fuller data DPK 30 Provinsi .....	86
Lampiran 5. Plot data Dana Pihak Ketiga setelah dilakukan Transformasi.....	87
Lampiran 6. Hasil Uji Augmented Dickey Fuller setelah dilakukan Transformasi.....	95
Lampiran 7. Jarak Penggabungan Tiap Stage Menggunakan Jarak Autocorrelation-based pada Analisis Cluster Hirarki..	96
Lampiran 8. Jarak Penggabungan Tiap Stage Menggunakan Jarak Complexity-Invariant pada Analisis Cluster Hirar.....	97
Lampiran 9. Jarak Penggabungan Tiap Stage Menggunakan Jarak Periodogram-based pada Analisis Cluster Hirarki .....	98
Lampiran 10. Jarak Penggabungan Tiap Stage Menggunakan Jarak Dynamic Time Warping pada Analisis Cluster Hirarki.....	99
Lampiran 11. Nilai Koefisien Silhouette pada Tiap Jumlah Kelompok yang Mungkin pada Analisis Cluster Menggunakan Jarak Dynamic Time Warping .....	100
Lampiran 12. Syntax Analisis Cluster Data Deret Waktu .....	101

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Analisis *cluster* merupakan metode pengelompokan multivariat (mempunyai banyak peubah) yang tujuan utamanya mengelompokan objek berdasarkan karakteristik yang dimiliki (Hair dkk, 2010). Analisis ini akan menghasilkan kelompok yang memiliki homogenitas tinggi di dalamnya dan heterogenitas tinggi antar kelompok. Analisis *cluster* mempunyai bidang aplikasi yang sangat luas, dan juga merupakan salah satu teknik multivariat yang banyak digunakan.

Analisis *cluster* bisa digunakan pada semua data tanpa memperhatikan skala pengukuran data tersebut (Liao, 2005). Namun analisis *cluster* sering juga digunakan pada data yang kompleks, yang membutuhkan teknik pengelompokan yang lebih mutakhir. Salah satunya contohnya analisis *cluster* pada data deret waktu.

Data deret waktu merupakan data yang diperoleh dari titik titik waktu, dimana terdapat struktur korelasi antara nilai-nilai pada data deret waktu tersebut (Box dkk, 2008). Oleh karena itu, prosedur dan teknik statistika untuk data independen tidak lagi dapat dilakukan, dan metode yang berbeda sangatlah dibutuhkan (Wei, 2006). Data deret waktu menarik untuk diteliti karena cakupan bidangnya yang luas, meliputi sains, bisnis, teknik, ekonomi, kesehatan, keuangan, sampai pemerintahan (Liao, 2005).

Analisis *cluster* pada data deret waktu tidak seperti analisis cluster pada data statis. Hal ini dikarenakan karakteristik data deret waktu yang dinamis, dengan struktur autokorelasi yang mendasar. Oleh karena itu analisis *cluster* pada data deret waktu perlu memperhatikan karakteristik data deret waktu yaitu nilai-nilainya berubah bergantung waktu (Manso, 2013). Metode pengelompokan konvensional umumnya memakai ukuran kedekatan yang statis karena diasumsikan bahwa kedekatan antar nilai observasi dilakukan pada satu waktu yang spesifik, tanpa mempertimbangkan hubungan saling ketergantungan antar nilai dalam observasi. Maka dalam analisis *cluster* pada data deret waktu, perlu ditentukan ukuran kedekatan yang dipakai, serta algoritma apa yang akan digunakan untuk mengelompokan data deret waktu (Liao, 2005).

Liao (2005) mengungkapkan bahwa terdapat beberapa algoritma yang telah dikembangkan untuk mengelompokkan data deret waktu. Algoritma-algoritma tersebut dimodifikasi dari algoritma yang biasa digunakan untuk data statis sehingga mampu digunakan pada data deret waktu atau dengan mengubah data deret waktu ke dalam bentuk data statis sehingga algoritma konvensional untuk data statis dapat diterapkan. Pendekatan yang pertama bekerja langsung pada data mentah deret waktu yang disebut pendekatan *raw-data based*. Langkah yang digunakan pada pendekatan ini adalah dengan mengganti jarak atau ukuran kesamaan untuk data statis dengan jarak yang lebih sesuai untuk data deret waktu. Pendekatan yang kedua, dilakukan dengan cara mengubah data mentah deret waktu ke dalam bentuk vektor dengan dimensi yang lebih rendah atau ke dalam bentuk parameter-parameter model, yang disebut pendekatan *feature-based* dan pendekatan *model-based*.

*Dynamic Time Warping* (DTW) adalah algoritma yang menghitung *warping path* yang optimal antara dua runtun data sehingga outputnya adalah nilai-nilai *warping path* dan jarak diantara kedua runtun tersebut (Sakoe dan Chiba,1978). Algoritma ini diperkenalkan sekitar tahun 60-an oleh Bellman dan Kalaba, baru sekitar tahun 70-an secara luas digunakan untuk aplikasi pengenalan suara (*speech recognition*) oleh Sakoe dan Chiba (1978).

Algoritma DTW disebut juga sebagai *non-linear sequence alignment*. *Sequence alignment* adalah cara yang mengatur atau menjajarkan sebuah deret dengan deret yang lain untuk mengidentifikasi kemiripan antar deret tersebut. Algoritma DTW lebih realistis untuk digunakan dalam mengukur kemiripan suatu pola (*pattern/template matching*) daripada hanya menggunakan algoritma pengukuran linier seperti Jarak *Euclidean*, *Manhattan*, *Canberra*, *Mexican Hat* dan lainnya. Data yang diolah selalu berada dalam domain waktu, sehingga urutan data yang kita punya dianggap bervariasi terhadap waktu (Sakoe dan Chiba, 1978).

Analisis *cluster* hirarki merupakan salah satu metode pengelompokan. Dalam metode hirarki data tidak dibagi menjadi kelas atau kelompok secara langsung dengan menentukan berapa jumlah *cluster* yang diinginkan. Sebaliknya klasifikasi terdiri dari serangkaian pembagian, yang diawali dari satu *cluster* yang berisi semua individu, menjadi  $n$  *cluster* masing-masing berisi satu individu (Everitt dkk,

2011). Analisis *cluster* hirarki memiliki beberapa metode penggabungan diantaranya : *single linkage* (pautan tunggal), *complete linkage* (pautan ), dan *average linkage* (pautan rata-rata). Ketiga jarak tersebut merupakan jarak yang paling sering digunakan, dan *average linkage* merupakan metode yang paling stabil dan tidak bias karena menggunakan rata-rata.(Everitt, 2011).

Analisis *cluster* hirarki memiliki kelemahan yaitu ketidakstabilan pengelompokan, perbedaan sedikit dalam input data akan menimbulkan hasil pengelompokan yang berbeda. Kelemahan lainnya adalah kesalahan dalam pengelompokan data yang memiliki pencilan. Pengelompokan hanya mengelompokan data pencilan dan yang bukan pencilan. (Hair dkk, 2010)

Algoritma *k-means* adalah salah satu metode pengelompokan secara non-hirarki. Metode pengelompokan non-hirarki kebalikan dari pengelompokan secara hirarki, jika pada pengelompokan hirarki jumlah kelompok tidak ditentukan terlebih dahulu, pada metode non-hirarki jumlah kelompok ditentukan terlebih dahulu, kemudian ditentukan pusat *cluster* masing-masing kelompok tersebut, untuk kemudian objek-objek dengan jarak terdekat dengan pusat *cluster* akan dikelompokkan menjadi satu kelompok (Hair dkk, 2010).

Analisis *cluster* menggunakan algoritma *k-means* memiliki banyak keunggulan, antara lain algoritma *k-means* mampu melakukan pengelompokan yang stabil, algoritma *k-means* dapat mengatasi permasalahan pencilan yang tidak dapat diatasi oleh analisis *cluster* secara hirarki, dan dapat mengelompokan data dengan jumlah yang besar (Hair dkk, 2010). Kelemahan Algoritma *k-means* , jika penentuan jumlah dan pusat *cluster* dilakukan secara acak hasil pengelompokan akan lebih buruk daripada analisis *cluster* hirarki.

Menurut Hair dkk (2010) banyak para ahli yang menyarankan agar mengkombinasikan metode pengelompokan hirarki dan metode pengelompokan non-hirarki untuk mendapatkan hasil yang lebih baik. Metode hirarki berfungsi untuk menentukan jumlah *cluster* optimal dan pusat masing-masing *cluster*, sedangkan metode non-hirarki berfungsi untuk mengatasi masalah pencilan pada metode hirarki serta menstabilkan hasil pengelompokan.

Aplikasi data deret waktu digunakan digunakan dalam berbagai bidang antara lain pada bidang ekonomi, finansial, kesehatan, dan bidang-bidang lainnya (Cryer dan Chan, 2008). Salah satu bidang yang

menarik untuk dipelajari adalah bidang ekonomi dan finansial, dimana data deret waktu digunakan untuk meramalkan suku bunga satu bulan ke depan, pendapatan tahunan, dan sebagainya (Wei, 2006). Salah satu data deret waktu dalam bidang ekonomi adalah dana nasabah bank, yang merupakan salah satu sumber dana yang digunakan oleh bank untuk menjalankan aktivitas ekonomi.

Dana Pihak Ketiga adalah dana yang berasal dari masyarakat (perorangan maupun badan usaha), dana ini dihimpun dalam bentuk giro, deposito dan tabungan. Dana Pihak Ketiga (DPK) merupakan sumber dana terbesar yang dimiliki oleh bank. DPK terbesar diperoleh dari nasabah bank umum, oleh karena itu bank umum memiliki peranan yang sangat besar dalam menggerakkan perekonomian di Indonesia (Dendawijaya, 2005)

Dalam penelitian ini akan dilakukan pengelompokan DPK dari seluruh bank umum di Indonesia. Data yang digunakan merupakan jumlah DPK perprovinsi selama Januari 2005 sampai dengan Desember 2013. Pengelompokan menggunakan jarak DTW, dengan metode pengelompokan secara hirarki, menggunakan pendekatan *average linkage* dengan optimalisasi menggunakan algoritma *k-means*.

## 1.2 Rumusan Masalah

1. Apakah jarak *Dynamic Time Warping* cocok untuk digunakan pada analisis *cluster* pada data deret waktu Dana Pihak Ketiga (DPK) di Indonesia Bulan Januari 2005 - Desember 2013?
2. Provinsi mana yang mempunyai pola pergerakan Dana Pihak Ketiga yang mirip ?
3. Bagaimana akurasi pengelompokan menggunakan jarak *Dynamic Time Warping* pada analisis *cluster* hirarki dengan optimalisasi pengelompokan menggunakan algoritma *k-means* ?

## 1.3 Tujuan Penelitian

1. Membandingkan hasil analisis *cluster* pada data deret waktu menggunakan jarak *Dynamic Time warping* dengan jarak pengelompokan deret waktu lainnya
2. Mengelompokan provinsi berdasarkan pola pergerakan DPK yang mirip

3. Mengetahui akurasi analisis *cluster* data deret waktu menggunakan jarak *Dynamic Time Warping* pada analisis *cluster* hirarki dengan optimalisasi menggunakan algoritma *k-means*

#### 1.4 Manfaat

Mengetahui penerapan jarak *Dynamic Time Warping* (DTW) pada data deret waktu dan kelompok yang dihasilkan serta akurasi pengelompokannya.

#### 1.5 Batasan Masalah

Analisis *cluster* menggunakan jarak DTW dan metode hirarki agglomeratif dengan optimalisasi menggunakan algoritma *k-means*. Data yang digunakan merupakan data sekunder jumlah Dana Pihak Ketiga dari seluruh provinsi di Indonesia pada Bulan Januari sampai dengan bulan Desember 2013. Untuk perhitungan kebaikan *cluster* menggunakan koefisien korelasi *cophenetic* pada analisis *cluster* hirarki dan menggunakan *silhouette width* pada algoritma *k-means*.



UNIVERSITAS BRAWIJAYA



## BAB II TINJAUAN PUSTAKA

### 2.1. Analisis *Cluster*

Analisis *Cluster* merupakan teknik dalam menemukan kelompok dalam kumpulan data. Pada dasarnya data dikelompokkan dengan tujuan data yang berada pada satu kelompok mempunyai kemiripan yang dekat dan memiliki perbedaan yang jelas dengan kelompok yang lain (Kaufman dan Rousseuw, 1990). Dengan kata lain secara statistika dapat dinyatakan bahwa data-data yang berada dalam satu kelompok memiliki homogenitas yang tinggi, dan antara satu kelompok dengan kelompok yang lain memiliki heterogenitas yang tinggi (Hair dkk, 2010).

Apabila pengelompokan ini sukses, individu-individu dalam kelompok akan menggerombol ketika diplot dan objek dalam kelompok lain akan terpisah jauh. Analisis *cluster* memiliki asumsi antar peubah saling bebas (tidak berkorelasi) dan data untuk seluruh peubah minimal memiliki skala interval (Fernandes, 2008).

Analisis *Cluster* dapat dikatakan sebagai analisis yang primitif karena tidak ada asumsi di dalamnya mengenai banyaknya kelompok yang terbentuk. Pengelompokan dilakukan atas dasar persamaan jarak.

#### 2.1.1 Analisis *Cluster* pada Data Statis

Han dan Kamber (2012) mengelompokkan metode pengelompokan pada data statis menjadi 5 macam yaitu metode hirarki, metode partisi atau metode non-hirarki, metode *density-based*, metode *grid-based*, dan metode *model-based*.

##### a. Metode partisi

Jika diberikan set data sejumlah  $n$ , metode partisi membuat  $k$  kelompok, dimana masing masing kelompok terdapat minimal satu objek dari data set dan  $k \leq n$ . Dua metode partisi yang sering sekali digunakan adalah algoritma *k-means*, dimana setiap kelompok diwakili oleh nilai rata-rata dari seluruh objek yang berada dalam kelompok tersebut. Algoritma yang kedua adalah algoritma *k-medoids* dimana setiap kelompok diwakili oleh objek yang menjadi pusat dari kelompok tersebut (Kaufman dan Rousseuw, 1990).

b. Metode Hirarki

Pengelompokan secara hirarki bekerja dengan mengelompokan data menjadi pohon kelompok. Ada dua cara yang umum dipakai pada metode hirarki : *agglomerative* (penggabungan) dan *devisive* (pemecahan). Metode penggabungan dimulai dengan menempatkan tiap objek sebagai kelompok tersendiri kemudian menggabungkan kelompok kelompok tersebut menjadi kelompok yang lebih besar. Proses ini dilakukan hingga semua objek berada pada satu kelompok atau sampai jumlah kelompok sudah sesuai dengan yang diinginkan. Metode pemecahan adalah kebalikan dari metode penggabungan. (Kaufman dan Rousseeuw, 1990)

c. Metode *Density-based*

Gagasan umum dari metode *Density-based* adalah kelompok terus berkembang selama densitas (jumlah objek atau data) masih berada dalam “*neighborhood*” pada beberapa batas tertentu. Salah satu algoritma yang paling sering digunakan adalah *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN), Algoritma ini membentuk area-area dengan kepadatan yang cukup tinggi ke dalam *cluster-cluster* dan menemukan *cluster-cluster* dalam bentuk yang sembarang dalam suatu *database* spasial yang memuat *noise* (Han dan Kamber, 2012).

d. Metode *Grid-based*

Metode *Grid-based* membatasi ruang objek menjadi sejumlah kotak yang membentuk struktur *grid* dimana semua operasi pengelompokan dilakukan. Salah satu contoh algoritma pada metode ini adalah algoritma STING, yang menggunakan beberapa kotak yang berbentuk persegi panjang dengan tingkat resolusi yang berbeda-beda. Kemudian informasi statistik pada tiap kotak dihitung dan disimpan. Proses pengelompokan dimulai pada struktur hirarki yang relatif rumit. Untuk setiap kotak dihitung selang kepercayaan yang menunjukkan relevansi dari tiap kotak pada tingkat resolusi yang berbeda. Kotak yang tidak relevan terus menerus disingkirkan sampai bagian terbawah resolusi didapat (Liao, 2005).

e. Metode *Model-based*

Metode *Model-based* memperkirakan model untuk tiap kelompok dan mencoba mencocokkan data kepada model terbaik yang diperkirakan. Ada dua pendekatan utama yang sering dipakai pada metode ini yaitu pendekatan secara statistik dan pendekatan dengan *neural network*. Contoh dari pendekatan secara statistik adalah algoritma *AutoClass* yang menggunakan analisis statistik bayesian untuk memperkirakan jumlah kelompok. Contoh pendekatan dengan *neural network* yang sudah dikenal adalah *Self Organizing Map* (SOM)

## 2.1.2 Analisis *Cluster* pada Data Deret Waktu

### 1. Data Deret Waktu

Deret waktu adalah sekelompok nilai pengamatan yang diperoleh pada titik waktu yang berbeda dengan selang waktu yang sama dan barisan data diasumsikan saling berhubungan satu sama lain. Model deret waktu adalah suatu model runtun waktu dimana observasi yang satu dengan yang lain saling berkorelasi (Box dan Jenkins, 2008). Menurut Cryer dan Chan (2008), deret waktu adalah serangkaian data pengamatan yang disusun menurut waktu, dimana data pengamatan tersebut bersifat acak dan saling berhubungan secara statistika.

Data deret waktu terdapat di berbagai bidang contohnya dalam bidang pertanian : produksi panen tahunan, dalam bidang ekonomi : rata-rata bunga mingguan, pendapatan tahunan, harga saham bulanan, dalam bidang teknik : pola suara, spektrum listrik, dalam bidang meteorologi : kecepatan angin perjam, dan suhu harian (Wei, 2006).

Menurut Wei (2006) sifat dasar dari deret waktu adalah nilai-nilai yang ada pada deret waktu berkaitan atau saling berkorelasi, sehingga urutan nilai-nilai tersebut menjadi sangat penting. Oleh karena itu, prosedur dan teknik statistika untuk data independen tidak lagi dapat dilakukan dan metode yang berbeda sangatlah dibutuhkan.

Plot deret waktu dapat memunculkan pola seperti pola acak, periodik atau siklis, trend, *level shifts*, pola tak biasa, serta gabungan dari beberapa pola (Montgomery, 2008).

### 2. Stasioneritas Deret Waktu

Menurut Wei (2006) deret waktu yang stasioner adalah relatif tidak terjadi kenaikan ataupun penurunan nilai secara tajam pada data atau fluktuasi data berada pada sekitar nilai rata-rata yang konstan. Data

secara kasarnya harus horizontal sepanjang sumbu waktu. Terdapat dua macam stasioneritas data yaitu stasioner terhadap ragam dan stasioner terhadap rata-rata.

a. Stasioneritas Terhadap Ragam

Data dikatakan stasioner terhadap ragam, apabila ragam dari data tidak berfluktuasi dari waktu ke waktu atau dapat dikatakan bahwa ragam konstan sepanjang waktu. Apabila data tidak stasioner terhadap ragam, maka data tersebut dapat ditransformasi dengan menggunakan transformasi Box Cox. Bentuk transformasi Box-Cox secara ringkas adalah sebagai berikut (Cryer dan Chan ,2008) :

$$T(Z_t) = Z_t^{(\lambda)} = \begin{cases} \frac{Z_t^{(\lambda)}}{\lambda}, & \lambda \neq 0 \\ \log(Z_t), & \lambda = 0 \end{cases} \quad (2.1)$$

dengan:

$T(Z_t)$  = data hasil transformasi

$Z_t$  = pengamatan pada waktu ke-t

$\lambda$  = parameter transformasi (diduga dengan metode kemungkinan maksimum)

b. Stasioneritas Terhadap Rata-rata

Data dikatakan stasioner terhadap rata-rata yakni data memiliki rata-rata yang konstan atau tidak terpengaruh oleh waktu pengamatan (Cryer dan Chan, 2008). Data yang stasioner terhadap rata-rata dapat diketahui dari plot autokorelasi, yaitu jika 95% plot autokorelasi dari data ( $\alpha = 0.05$ ) masuk ke dalam selang  $\pm 1,96(1/\sqrt{n})$  atau jika *Autocorrelation Function* berbeda nyata pada  $\text{lag} \leq 3$ .

Data yang terindikasi adanya *trend* dikatakan tidak stasioner terhadap rata-ratanya. Jika tidak ada variasi *trend* maka deret waktu sudah stasioner terhadap rata-rata. Makridakis dkk (1999) mengatakan bahwa pada data yang stasioner, nilai-nilai autokorelasi akan turun sampai nol sesudah *time lag* kedua atau ketiga sedangkan untuk data yang tidak stasioner, nilai-nilai tersebut berbeda signifikan dari nol untuk beberapa periode waktu.

Pemeriksaan stasioneritas pada rata-rata dapat menggunakan uji *Augmented Dickey Fuller (ADF)* .). Untuk memudahkan pemahaman mengenai uji *Augmented Dickey Fuller*, diberikan model AR (1) :

$$Z_t = \Phi Z_{t-1} + e \quad (2.2)$$

Jika persamaan (2.4) dikurangi  $Z_{t-1}$  pada sisi kanan dan kiri, maka :

$$Z_t - Z_{t-1} = \Phi Z_{t-1} - Z_{t-1} + e_t \quad (2.3)$$

$$\Delta Z_t = \{\Phi - 1\} Z_{t-1} + e_t$$

$$\Delta Z_t = \Phi^* Z_{t-1} + e_t$$

di mana  $e_t \sim IID(0, \sigma^2)$ .

Dari persamaan (2.5) dapat dibuat hipotesis :

$H_0 : \Phi^* = 0$  (data tidak stasioner)

$H_1 : \Phi^* < 0$  (data stasioner)

Statistik uji yang digunakan adalah :

$$\tau = \frac{\hat{\Phi}^*}{SE(\hat{\Phi}^*)} \quad (2.4)$$

di mana

$\hat{\Phi}^*$  = nilai duga parameter *autoregressive* (AR)

$SE(\hat{\Phi}^*)$  = *standard error*  $\hat{\Phi}^*$

$$SE(\hat{\Phi}^*) = \frac{\sigma(\hat{\Phi}^*)}{\sqrt{n}} \quad (2.5)$$

dengan  $\sigma(\hat{\Phi}^*)$  adalah simpangan baku dari  $\hat{\Phi}^*$  dan  $n$  adalah banyaknya pengamatan. Nilai statistic uji  $\tau$  dibandingkan dengan tabel Dickey Fuller dengan tingkat kesalahan sebesar  $\alpha$  dan derajat bebas  $n$ ,  $\tau \sim \tau_{(a,n)}$  berdasarkan kriteria uji sebagai berikut :

- Tidak ada cukup bukti untuk menerima  $H_0$  atau data tidak stasioner jika  $\tau > \tau_{(a,n)}$
- Cukup bukti untuk menerima  $H_0$  atau data stasioner jika  $\tau \leq \tau_{(a,n)}$

Pada  $\alpha = 0.05$ , jika *p-value* kurang dari  $\alpha$  maka  $H_0$  ditolak, artinya deret stasioner, sedangkan penerimaan  $H_0$  menunjukkan bahwa deret tidak stasioner, sehingga perlu dilakukan differencing (Harris dan Sollis, 2003).

Jika suatu data deret waktu tidak stasioner terhadap rata-rata, maka data dapat diubah menjadi deret stasioner melalui *differencing* (pembedaan). Banyaknya *differencing* yang dilakukan untuk mencapai keadaan stasioner dinotasikan dengan  $d$ . Proses *differencing* ini bisa dilakukan hingga orde ke- $d$ . Jika proses *differencing* dilakukan sebanyak satu kali dinamakan *differencing* dengan orde  $d = 1$ , bila dilakukan *differencing* sebanyak dua kali maka dinamakan *differencing* dengan *orded* = 2 dan seterusnya. *Differencing* pertama

( $d = 1$ ) dari  $X_t$  dapat dituliskan sebagai berikut (Cryer dan Chan, 2008) :

$$\nabla Y_t = Y_t - Y_{t-1} \quad (2.6)$$

dimana:

$Y_t$  = nilai pengamatan pada waktu ke  $t$

$Y_{t-1}$  = nilai pengamatan pada waktu ke  $t - 1$

$\nabla Y_t$  = hasil *differencing* pertama pengamatan  $Y_t$

sedangkan bentuk *differencing* kedua ( $d = 2$ ) adalah :

$$\nabla^2 Y_t = \nabla Y_t - \nabla Y_{t-1} \quad (2.7)$$

di mana :

$Y_t$  = pengamatan pada periode waktu ke -  $t$

$Y_{t-1}$  = pengamatan pada periode waktu ke - ( $t-1$ )

$\nabla Y_t$  = hasil *differencing* pertama pengamatan  $Y_t$

$\nabla Y_{t-1}$  = hasil *differencing* pertama pengamatan  $Y_{t-1}$

$\nabla^2 Y_t$  = hasil *differencing* kedua pengamatan  $Y_t$

*Differencing* dilakukan hingga data hasil *differencing* menunjukkan keadaan yang stasioner terhadap rata-rata.

### 3. Metode Pengelompokan

Sama halnya dengan analisis *cluster* pada data statis, analisis *cluster* pada data deret waktu membutuhkan algoritma pengelompokan atau prosedur untuk membentuk kelompok-kelompok yang didapat dari dataset, dan pemilihan algoritma pengelompokan berdasarkan dari tipe data yang akan dikelompokkan dan tujuan tertentu dari peneliti. Selama berkembangnya pengelompokan pada data deret waktu, banyak teknik dikembangkan baik itu untuk data uniform atau non-uniform, univariat atau multivariat, dan untuk data dengan panjang yang sama maupun tidak (Liao,2005).

Berbagai-macam algoritma telah dikembangkan untuk mengelompokan data deret waktu, meskipun ada berbagai macam metode dalam mengelompokan data deret waktu, dapat dikatakan bahwa semua algoritma tersebut mencoba memodifikasi algoritma yang telah digunakan pada data statis sehingga mampu mengatasi pengelompokan data multivariat atau dapat mengubah data deret waktu ke dalam bentuk data statis sehingga algoritma untuk data statis dapat langsung digunakan (Liao, 2005).

Ada tiga pendekatan yang telah dilakukan untuk melakukan data deret waktu, antara lain *raw-data-based*, *feature-based*, dan *model-based*.



Gambar 2.1. Tiga pendekatan pengelompokan data deret waktu (a) *raw-data-based*, (b) *feature-based*, (c) *model-based*

a. *Raw-data-based*

Pendekatan ini biasanya digunakan secara langsung terhadap data deret waktu, modifikasi pada metode ini adalah dengan

menggantikan jarak atau ukuran kesamaan untuk data statis dengan jarak atau ukuran yang cocok untuk data deret waktu. (Sangeeta dan Geeta,2012)

Beberapa jarak atau ukuran kesamaan telah diperkenalkan untuk mengatasi pengelompokan data deret waktu yang saling berkorelasi dan bergerak dinamis terhadap waktu antara lain : jarak *Pearson's Correlation Coefficient*, *Short Time Series*, *Kullback-liebre*, *Cross Correlation Function* antara dua deret waktu (Liao, 2005).

#### b. *Featured-based*

Pendekatan ini bermula karena data mentah yang didapat selalu dalam dimensi yang tinggi (*high-dimensional*) sehingga ditakutkan terjadi kesalahan dalam proses pengelompokan, terutama pengelompokan dengan kecepatan tinggi, seperti menggunakan *Self Organizing Map* (SOM). Pendekatan *feature-based* membutuhkan *Feature Extraction* yang merupakan suatu pengambilan ciri atau fitur dari suatu bentuk, dan dapat dikatakan sebagai karakteristik dari suatu objek yang dapat dibedakan oleh bentuk, tekstur, ukuran, intensitas, sifat statistik, dan sebagainya (Liao, 2005).

Fitur yang akan diidentifikasi harus membawa informasi yang dirangkum dari deret waktu, yang menangkap gambaran global dari data. Fitur dari trend, musiman, periodik, korelasi serial, kemiringan, dan kurtosis telah banyak digunakan sebagai langkah-langkah dalam banyak penelitian pada pengelompokan data deret waktu *Feature-based*.

Berbagai macam metode *feature extraction* telah dikembangkan, antara lain, *Hear Wavelet Transform*, *Independent Component Analysis*, *Wavelet Transform*, *Hough Transform*, *Discrete Time Fourier Transform*, dan *Principal Component Analysis*. Masing masing metode tersebut bekerja dengan melakukan ekstraksi fitur terhadap data mentah berupa deret waktu, dan output yang diberikan berupa hasil ekstraksi yang telah mereduksi dimensi yang tinggi pada deret waktu, sehingga analisis *cluster* pada data statis bisa dapat dilakukan pada data ekstraksi tersebut (Sangeeta, 2012).

#### c. *Model-based*

Pendekatan ini menganggap setiap deret waktu dihasilkan oleh beberapa macam model atau campuran dari distribusi distribusi probabilitas yang mendasar. Deret waktu dikatakan sama atau mirip jika karakteristik individual model atau residual setelah menyesuaikan model itu sama atau mirip (Liao,2005).

Untuk mengelompokkan atau memilih model dari sekelompok struktur dinamis (khususnya pada model model invertibel ARIMA) Picollo (1990) memperkenalkan jarak *euclidean* antar perluasan nilai *autoregressive* yang bersesuaian sebagai metrik. Metrik memenuhi kriteria klasik dari jarak, yaitu, non-negatifitas, simetrisitas, dan triangularitas. Matriks jarak di antara sepasang model deret waktu kemudian di kelompokkan menggunakan metode *complete linkage* untuk membuat dendrogram.

Maharaj (2000) mengembangkan prosedur pengelompokan secara hirarki (*agglomerative*) berdasar pada uji hipotesis menggunakan *p-value* diterapkan untuk setiap pasang dari deret waktu stasioner yang telah diberikan. Asumsi pada prosedur ini adalah bahwa setiap deret waktu stasioner dapat dicocokkan dengan model linier  $AR_{(k)}$  dinotasikan dengan parameter parameter vektor  $\pi' = [\pi_1, \pi_2, \dots, \pi_k]$ . Uji statistik distribusi *chi-square* digunakan untuk uji *null hipotesis* ada atau tidaknya perbedaan antara dua proses deret waktu yang dihasilkan atau dapat dituliskan  $H_0: \pi_x = \pi_y$ . Dua deret waktu dikelompokkan bersama apabila nilai *p-value* lebih besar daripada *pre-spesified significance level*. masih terdapat banyak lagi pengelompokan dengan pendekatan *model-based* lain yang masih dikembangkan oleh para ilmuwan statistika.

## 2.2. Ukuran Kemiripan

### 2.2.1. Jarak Euclidian

Persamaan yang umum digunakan untuk menghitung jarak antara objek  $j$  dengan objek  $k$  adalah sebagai berikut:

$$d(j, k) = \sqrt{\sum_{i=1}^n (X_{ij} - X_{ik})^2} \quad (2.8)$$

dengan  $X_{ij}$  dan  $X_{ik}$  adalah nilai dari peubah ke- $i$  objek ke- $j$  dan objek- $k$  dan  $p$  adalah banyaknya peubah (Marrison, 1978).

Persamaan 2.8 disebut juga sebagai jarak *euclidean*. Penghitungan jarak tersebut umum digunakan karena dapat menghitung jarak geometrik dalam ruang dimensi ganda. Sebelum dilakukan penghitungan jarak dengan metode jarak *euclidean*, peubah yang akan dianalisis harus memenuhi 3 syarat, yaitu: peubah tidak saling berkorelasi, memiliki satuan pengukuran yang sama, dan pengukuran terstandarisasi (Manly, 1988).

Marrison (1978) menjelaskan bahwa jarak *Euclidean* tergantung skor baku, yaitu dengan mentransformasi peubah asal sehingga memiliki nilai tengah 0 dan simpangan baku 1. Model transformasinya adalah sebagai berikut:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_i}{S_i} \quad (2.9)$$

di mana:  $X_{ij}$  = pengamatan peubah ke-i item ke-j

$\bar{X}_i$  = rata-rata peubah ke-i

$S_i$  = simpangan baku peubah ke-i

Sehingga jarak *Euclidean* antara item ke-j dengan item ke-k dari transformasi di atas adalah

$$d(Z_{ij}, Z_{ik}) = \sqrt{\sum_{i=1}^p (Z_{ij} - Z_{ik})^2} \quad (2.10)$$

Apabila terdapat banyak peubah, maka hasil perhitungan jarak dapat diringkas dalam bentuk matriks jarak (D) ukuran n x n, dengan n adalah banyaknya item yang diamati.

$$D = \begin{bmatrix} 0 & d_{12} & d_{13} & \dots & d_{1n} \\ d_{21} & 0 & d_{23} & \dots & d_{2n} \\ d_{31} & d_{32} & 0 & \dots & d_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & d_{n3} & \dots & 0 \end{bmatrix} \quad (2.11)$$

### 2.2.2. *Periodogram-Based Distance*

Caido dkk (2006) memperkenalkan beberapan jarak berdasarkan periodogram pada data deret waktu. Misalkan  $I_X(\lambda_k) = T^{-1}|\sum_{t=1}^T X_t e^{i\lambda_k t}|^2$  dan  $I_Y(\lambda_k) = T^{-1}|\sum_{t=1}^T Y_t e^{i\lambda_k t}|^2$  adalah periodogram dari dua deret waktu X dan Y, pada frekuensi  $\lambda_k = 2\pi k/T$ ,  $k=1,2,\dots,n$ , dengan  $n = [(T-1)/2]$ . Berdasarkan kedua periodogram tersebut beberapa jarak didefinisikan sebagai berikut.

Jarak *euclidean* antara dua ordinat periodogram:

$$d_P(X_t, Y_t) = \frac{1}{n} \left\{ \sum_{k=1}^n (I_X(\lambda_k) - I_Y(\lambda_k)) \right\}^{1/2} \quad (2.12)$$

Apabila data deret waktu yang didapat mempunyai satuan yang berbeda, hasil yang lebih baik akan didapatkan menggunakan jarak *euclidean* antara dua ordinat periodogram yang telah dibakukan

$$d_{NP}(X_t, Y_t) = \frac{1}{n} \left\{ \sum_{k=1}^n (NI_X(\lambda_k) - NI_Y(\lambda_k)) \right\}^{1/2} \quad (2.13)$$

di mana  $NI_X(\lambda_k) = I_X(\lambda_k) / \hat{\gamma}_0^X$  dan  $NI_Y(\lambda_k) = I_Y(\lambda_k) / \hat{\gamma}_0^Y$ , di mana

$\hat{\gamma}_0^X$  dan  $\hat{\gamma}_0^Y$  adalah ragam sampel dari masing masing deret X dan Y.

Karena ragam dari ordinat periodogram cocok untuk nilai spektrum pada frekuensi yang saling berhubungan, maka kita juga dapat menggunakan logaritma dari periodogram yang telah dibakukan

$$d_{LNP}(X_t, Y_t) = \frac{1}{n} \left\{ \sum_{k=1}^n (\log NI_X(\lambda_k) - \log NI_Y(\lambda_k)) \right\}^{1/2} \quad (2.14)$$

Casado (2010) mengungkapkan suatu jarak berdasarkan versi kumulatif dari periodogram-periodogram, yaitu periodogram terintegrasi. Pendekatan berdasarkan periodogram terintegrasi menunjukkan beberapa keuntungan dari pada jarak jarak berdasarkan periodogram lainnya, antara lain :

1. Mempunyai unsur-unsur asimtotik yang baik. Periodogram adalah penduga tak bias secara asimtotik namun tidak konsisten sebagai penduga pada densitas spektral, sedangkan periodogram terintegrasi adalah penduga yang konsisten pada distribusi spektral.
2. Dalam sudut pandang secara teori, distribusi spektral selalu dijumpai, densitas spektral tentu hanya ada dibawah distribusi kontinu. Meskipun dalam praktiknya, spektrum terintegrasi biasanya diduga dengan penduga spektrum.
3. Periodogram terintegrasi dapat menjelaskan secara menyeluruh proses stokastik

Casado (2010) mengemukakan dua jenis jarak berdasarkan periodogram terintegrasi, yaitu yang dinormalkan dan yang tidak dinormalkan. Periodogram yang dinormalkan lebih memperhatikan pada bentuk kurva, sedangkan yang tidak dibakukan lebih memperhatikan skala. Casado (2010) menyarankan agar menggunakan periodogram

yang dinormalkan ketika grafik fungsi cenderung saling berpotongan, sedangkan yang tidak dinormalkan jika grafik fungsi tidak saling berpotongan. Jarak berdasarkan periodogram terintegrasi memiliki persamaan :

$$d_{IP}(X_t, Y_t) = \int_{-\pi}^{\pi} |F_X(\lambda) - F_Y(\lambda)| d\lambda \quad (2.15)$$

di mana  $F_X(\lambda_j) = C_X^{-1} \sum_{i=1}^j I_X(\lambda_i)$  dan  $F_Y(\lambda_j) = C_Y^{-1} \sum_{i=1}^j I_Y(\lambda_i)$ , dengan  $C_X = \sum_i I_X(\lambda_i)$  dan  $C_Y = \sum_i I_Y(\lambda_i)$  untuk periodogram terintegrasi yang dinormalkan, dan  $C_X=1$  dan  $C_Y=1$  untuk yang tidak dinormalkan.

### 2.2.3. Complexity-invariance distance

Jarak ini menggunakan informasi tentang perbedaan kompleksitas antara dua deret waktu sebagai faktor koreksi dari jarak yang telah ada. Dalam hal ini jarak yang sering digunakan adalah jarak *euclidean*. Kompleks dalam deret waktu biasanya sering diartikan sebagai banyaknya puncak dan nilai yang berubah menurut waktu. Jadi jarak ini menyempurnakan jarak yang telah digunakan untuk data statis yang memiliki kelemahan jika digunakan pada deret waktu, karena strukturnya yang berbeda dengan data statis. (Batista dkk, 2011)

Jarak *complexity-invariance* dapat dirumuskan sebagai faktor koreksi dari jarak *euclidean* sebagai berikut :

$$CID(Q, C) = ED(Q, C) \times CF(Q, C) \quad (2.16)$$

di mana  $ED(Q, C)$  adalah jarak *euclidean* dari dua deret waktu  $Q$  dan  $C$ , sedangkan  $CF(Q, C)$  adalah faktor koreksi kompleksitas yang didefinisikan sebagai:

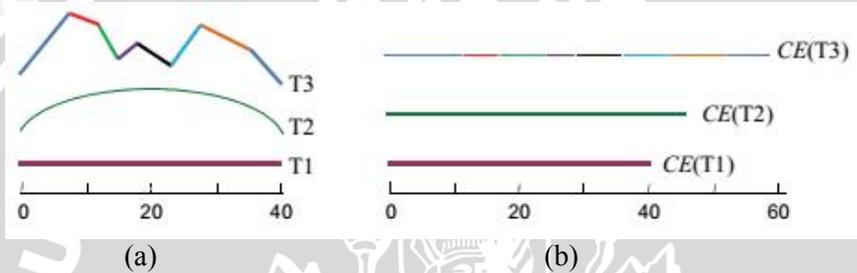
$$CF(Q, C) = \frac{\max(CE(Q), CE(C))}{\min(CE(Q), CE(C))} \quad (2.17)$$

dan  $CE(T)$  adalah estimasi kompleksitas dari deret waktu  $T$ .

Sebelum dijelaskan bagaimana penggunaan  $CE$ , ada baiknya memperhatikan bahwa  $CF$  memperhitungkan perbedaan kompleksitas pada deret waktu yang dibandingkan.  $CF$  memaksa deret waktu dengan berbagai perbedaan kompleksitas agar dapat dianalisis lebih lanjut. Dalam kasus ini semua deret waktu memiliki kompleksitas yang sama,

CID dapat digunakan pada jarak *euclidean* secara mudah (Batista, dkk 2011).

Pendekatan yang dilakukan oleh Batista dkk (2011), dalam mengestimasi kompleksitas dalam deret waktu sangatlah sederhana. Pendekatan ini berdasarkan pada intuisi fisik yang dapat disebut *stretch* pada deret waktu hingga menjadi garis lurus, deret waktu kompleks akan menghasilkan garis yang lebih panjang dari pada deret waktu sederhana. Sebagai contoh diberikan ilustrasi seperti berikut



Gambar 3.1. (a) pola deret waktu dengan kompleksitas yang berbeda (b) nilai kompleksitas deret waktu berdasarkan lengkungannya.

Estimasi kompleksitas dapat dihitung dengan persamaan 2.18 berikut

$$CE(Q) = \sqrt[2]{\sum_{i=1}^{n-1} (q_i - q_{i-1})^2} \tag{2.18}$$

di mana  $q_i$  adalah nilai ke  $i$  dari deret waktu  $q$

Pendekatan CID untuk estimasi kompleksitas dapat secara mudah digunakan pada bahasa pemrograman apapun.

#### 2.2.4. Correlation-based Distance

Jarak ini didasarkan pada faktor korelasi *pearson* antara dua deret waktu. Misalkan  $x_i$  dan  $v_j$  masing masing adalah vektor dengan  $P$  dimensi. Faktor korelasi *pearson* antara  $x_i$  dan  $v_j$ ,  $cc$  didefinisikan sebagai:

$$cc = \frac{\sum_{k=1}^P (x_{ik} - \mu_{ik})(v_{jk} - \mu_{jk})}{S_{x_i} S_{v_j}} \tag{2.19}$$

di mana  $\mu_{x_i}$  dan  $S_{x_i}$  masing masing adalah rata-rata dan simpangan baku dari  $x_i$ , yang dihitung seperti berikut:

$$\mu_{x_i} = \frac{1}{p} \sum_{k=1}^p x_{ik} \quad (2.20)$$

$$S_{x_i} = \left[ \sum_{k=1}^p (x_{ik} - \mu_{x_i})^2 \right]^{\frac{1}{2}} \quad (2.21)$$

Dua jarak *cross-correlation-based* yang digunakan oleh pada algoritma *fuzzy c-means* adalah:

$$d_{cc}^1 = \left( \frac{1-cc}{1+cc} \right)^\beta \quad (2.22)$$

dan

$$d_{cc}^2 = 2(1 - cc) \quad (2.23)$$

Pada persamaan 2.15,  $\beta$  mempunyai fungsi yang sama dengan  $m$  pada algoritma *fuzzy c-means* yaitu bilangan real yang lebih besar dari 0 yang digunakan sebagai aturan pada algoritma *fuziness* (Manso, 2013).

### 2.2.5. Short Time Series Distance

Jarak ini digunakan sebagai ukuran kesamaan untuk data deret waktu *microarray*. Ukuran kemiripan ini didasarkan pada konsep kemiripan dan karakteristik tertentu dari deret waktu yang dihasilkan pada eksperimen *microarray*. *Microarray* merupakan suatu pola yang didapatkan dari analisa fungsi dan ekspresi gen dalam jumlah yang banyak secara simultan dan dalam satu kali percobaan saja. Perlu diketahui tidak ada definisi yang jelas tentang dua deret waktu dapat dikatakan “mirip” dalam konteks biologis. Namun, secara umum dipahami bahwa dua deret waktu dapat dikatakan mirip jika bentuk deret waktu tersebut mirip. Deret *microarray* ini memiliki dua sifat utama yang ditentukan oleh sifat eksperimen yang menghasilkan deret *microarray* tersebut deret tersebut biasanya pendek dan pengambilan sampelnya tidak merata. Pada deret waktu yang pendek analisis statistik tradisional tidak selalu cocok, sehingga dibutuhkan ukuran yang sesuai untuk mengukur kemiripan antara dua deret waktu yang pendek tersebut, yaitu dengan menganggap setiap deret waktu sebagai potongan dari sebuah fungsi linier. Jarak *Short time series* (STS) dinyatakan sebagai jumlah kuadrat selisih slope antara dua deret waktu. Jarak STS antara dua deret waktu  $x_i$  dan  $v_j$  adalah:

$$d_{STS} = \sqrt{\sum_{k=1}^p \left( \frac{v_{j(k+i)} - v_{jk}}{t_{(k+1)} - t_k} - \frac{x_{i(k+1)} - x_{ik}}{t_{(k+1)} - t_k} \right)^2} \quad (2.24)$$

Di mana  $t_k$  adalah waktu dari tiap titik pada data  $x_{ik}$  dan  $v_{jk}$ . Untuk menghilangkan efek dari skala, standarisasi pada deret waktu sangat dianjurkan (Moler dkk, 2003).

### 2.2.6. Autocorrelation-based distance

Galeano dan Pena (2000) mengusulkan sebuah metrik berdasarkan pendugaan *autocorrelation function* (ACF) untuk data deret waktu yang mempunyai struktur korelasi yang menarik untuk diteliti.

Misalkan  $\hat{\rho}_X = (\hat{\rho}_{1,X}, \dots, \hat{\rho}_{L,X})^t$  dan  $\hat{\rho}_Y = (\hat{\rho}_{1,Y}, \dots, \hat{\rho}_{L,Y})^t$  adalah vektor-vektor autokorelasi hasil pendugaan dari deret waktu X dan Y, untuk beberapa L seperti  $\hat{\rho}_{i,X} \approx 0$  dan  $\hat{\rho}_{i,Y} \approx 0$  untuk  $i > L$ . Jarak antara dua deret waktu kemudian dapat dibentuk sebagai berikut :

$$d_{ACF} = \{(\hat{\rho}_X - \hat{\rho}_Y)^t \Omega (\hat{\rho}_X - \hat{\rho}_Y)\}^{1/2} \quad (2.25)$$

di mana  $\Omega$  adalah matriks pembobot. (Manso, 2013) Beberapa macam bentuk persamaan yang dihasilkan dengan  $\Omega$  berbeda sebagai berikut :

- Apabila pembobot disamakan dengan menjadikan  $\Omega = I$ . Pada beberapa kasus  $d_{ACF}$  menjadi jarak *euclidean* antara dua fungsi autokorelasi yang diduga

$$d_{ACFU}(X_t, Y_t) = \left\{ \sum_{i=1}^L (\hat{\rho}_X - \hat{\rho}_Y)^2 \right\}^{1/2} \quad (2.26)$$

- Mempertimbangkan bobot geometris tidak berfungsi pada lag autokorelasi, sehingga  $d_{ACF}$  mempunyai bentuk persamaan:

$$d_{ACFG} = \left\{ \sum_{i=1}^L (p(1-p)^i) (\hat{\rho}_X - \hat{\rho}_Y)^2 \right\}^{1/2} \quad (2.27)$$

dengan  $0 < p < 1$

- Apabila  $\Omega = \text{Cov}(\hat{\rho})^{-1}$ , invers matriks kovarian dari autokorelasi, maka jarak yang dibentuk adalah jarak Mahalanobis antar autokorelasi, dapat ditulis sebagai  $d_{ACFM}$

Semua ukuran kemiripan tersebut sampai sekarang digunakan pada domain waktu, namun pendekatan pada domain frekuensi juga menjadi alternatif yang menarik untuk mengukur kemiripan antar deret waktu (Manso, 2013).

### 2.3. Jarak Dynamic Time Warping

*Dynamic Time Warping* (DTW) pertama kali diperkenalkan pada tahun 1960-an oleh R. Bellman, dan di eksplorasi secara luas pada tahun 1970-an untuk aplikasi pengenalan suara oleh C. Myers. Saat ini, metode *Dynamic Time Warping* sudah diterapkan dalam berbagai bidang, diantaranya: *handwriting* dan *online signature matching* oleh C.C. Tappert pada tahun 1990, *computer vision* dan *computer animation* oleh Muller pada tahun 2007, *protein sequence alignment* dan *chemical engineering* oleh J. Vial pada tahun 2008, dan lain – lain (Senin, 2008).

Menurut Sakoe dan Chiba (1978), *dynamic time warping* (DTW) adalah metode untuk menghitung kesamaan antara dua deret waktu yang mungkin berbeda dalam waktu dan kecepatan. Contoh, dalam kasus mendeteksi kesamaan pola gerakan berjalan, dimana data yang pertama menunjukkan seseorang berjalan secara perlahan-lahan dan data lainnya menunjukkan seseorang tersebut berjalan lebih cepat.

DTW disebut juga sebagai *non-linear sequence alignment*, sehingga jarak ini lebih realistis untuk digunakan dalam mengukur kemiripan suatu pola (*pattern/template matching*) daripada hanya menggunakan algoritma pengukuran linier seperti jarak *Euclidean*, *Manhattan*, *Canberra*, *Mexican Hat* dan lainnya. Data yang diolah selalu berada dalam domain waktu, sehingga rentetan data yang kita punya dianggap bervariasi terhadap waktu. *Dynamic Time Warping* (DTW) adalah generalisasi dari algoritma klasik untuk membandingkan urutan diskrit dengan urutan nilai nilai kontinu. *Dynamic Time Warping* mengatasi keterbatasan jarak *euclidian* yang hanya bisa digunakan pada deret waktu pada selang waktu yang sama, dan juga dapat digunakan pada deret waktu dengan panjang yang tidak sama (Niennattrakul dan Ratanamahatana, 2007).

*Dynamic Time Warping* menggunakan teknik pemrograman dinamis untuk menemukan semua jalur yang mungkin, dan memilih salah satu yang menghasilkan jarak minimum antara dua deret waktu menggunakan matriks jarak, dimana masing-masing elemen dalam matriks adalah jarak kumulatif nilai minimal dari tiga tetangga sekitarnya. Misalkan terdapat dua deret waktu yaitu  $Q = q_1, q_2, \dots, q_i, \dots, q_n$  and  $C = c_1, c_2, \dots, c_j, \dots, c_m$ , pertama kita buat matriks berukuran  $n \times m$ , untuk setiap elemen  $(i,j)$  dalam matriks yang merupakan jarak kumulatif dari jarak  $(i,j)$  dan nilai minimum dari tiga

elemen yang berdekatan dengan elemen  $(i,j)$ , dimana  $0 < i \leq n$  dan  $0 < j \leq n$ , kita dapat mendefinisikan elemen  $(i,j)$  sebagai :

$$e_{ij} = d_{ij} + \min \{e_{(i-1)(j-1)}, e_{(i-1)j}, e_{i(j-1)}\} \quad (2.28)$$

Di mana  $d_{ij} = (q_i - c_j)^2$  dan  $e_{ij}$  adalah elemen  $(i,j)$  dari matriks yang merupakan jumlah akhir kuadrat jarak antara  $q_i$  dan  $c_j$ , dan jarak kumulatif terkecil dari tiga elemen di sekitar elemen  $(i,j)$ . Kemudian, untuk menemukan *optimal path*, kita harus menentukan *path* yang memberikan jarak kumulatif terkecil pada  $(n,m)$ . Jarak kumulatif terkecil pada  $(n,m)$  didefinisikan sebagai :

$$D_{DTW}(Q, C) = \min_{w \in P} \left\{ \sqrt{\sum_{k=1}^K d_{w_k}} \right\} \quad (2.29)$$

Di mana  $P$  adalah sekumpulan dari semua *warping path* yang mungkin, dan  $w_k$  adalah elemen  $(i,j)$  ke  $k$  pada *warping path* dan  $K$  adalah panjang dari *warping path* (Niennattrakul dan Ratanamahatana, 2007). Senin (2008) menguraikan algoritma dan kriteria yang untuk menghitung *optimal path warping*:

#### 1. Menentukan matriks *cost*

Metode ini diawali dengan menentukan matriks *cost* yang memiliki dimensi *cost*  $(n \times m)$ . Matriks *cost* adalah matriks yang merepresentasikan selisih dari Himpunan  $X$  dan  $Y$ . Masing-masing elemen matriks dapat dihitung dengan persamaan berikut :

$$C_p = \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} & \dots & c_{1,j} \\ c_{2,1} & c_{2,2} & c_{2,3} & \dots & c_{2,j} \\ c_{3,1} & c_{3,2} & c_{3,3} & \dots & c_{3,j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{i,1} & c_{i,2} & c_{i,3} & \dots & c_{i,j} \end{bmatrix} \quad (2.30)$$

Di mana :

$$C_p \in R^{n \times m}: c_{i,j} = \|x_i - y_j\|, i \in [1:n], j \in [1:m] \quad (2.31)$$

Nilai dari *cost* setiap matriksnya, digunakan untuk mencari jalur dengan *cost* terendah dari kolom pertama sampai kolom terakhir.

#### 2. Menentukan *optimal warping path*

Setelah matriks *cost* terbentuk, algoritma ini akan mencari jalur *warping* dengan *cost* terendah  $p = \{p_1, p_2, \dots, p_k\}$  dengan  $p = (p_i, p_j) \in [1:n] \times [1:m]$

Dengan Jalur *warping* yang akan dicari harus memenuhi kriteria-kriteria sebagai berikut:

a. *Monotonicity*

Artinya proses mengikuti urutan berdasar waktu.  $n_1 \leq n_2 \leq \dots \leq n_k$  dan  $m_1 \leq m_2 \leq \dots \leq m_k$ . Tujuannya agar proses bisa dilakukan straightforward sehingga kita tidak mengolah data yang sama berulang kali (tidak terjadi perulangan).

b. *Boundary Condition*

Artinya titik awal dan titik akhir dari jalur *warping* adalah titik awal dan titik akhir dari sekuen/rangkaian data (*sequence*). Ini adalah sebuah keharusan.  $p_1=(1,1)$  dan  $p_k=(m,n)$  Tujuannya data yang kita punya diproses seluruhnya mulai dari awal hingga akhir.

c. *continuity*

Artinya indeks dari data yang akan diproses selanjutnya tidak serta-merta melompat ke indeks yang jaraknya jauh. Jarak antara data yang pertama dengan yang selanjutnya adalah maksimal 1.  $Indeksdata2 - indeksdata1 \leq 1$ . Tujuannya agar tidak ada bagian dari data yang memang penting yang tidak diikuti dalam proses.

3. Menentukan matriks DTW

Perhitungan untuk matriks DTW adalah sebagai berikut:

a. Baris pertama:

$$D(1, j) = \sum_{k=1}^j c(x1, yk), j \in [1, M] \quad (2.32)$$

b. Kolom pertama:

$$D(i, 1) = \sum_{k=1}^i c(xk, y1), i \in [1, N] \quad (2.33)$$

c. Semua elemen matriks:

$$D(i, j) = c(x_i, y_j) + \min\{D(i - 1, j - 1), D(i - 1, j), D(i, j - 1), i \in 1, N, j \in 1, M\} \quad (2.34)$$

4. Menentukan *cost optimal warping path*

Setelah menghitung matriks DTW, maka untuk menentukan total dari *cost* jarak yang melewati jalur *warping* dapat menggunakan persamaan berikut:

$$cost = \frac{1}{m} \{ \min[DTW(1, m), DTW(2, m), \dots, DTW(n, m)] \} \quad (2.35)$$

Dengan *cost* dari *optimal warping path* adalah *cost* akumulasi minimum pada kolom terakhir.

## 2.4. Metode Hirarki

Metode hirarki merupakan metode pengelompokan yang terstruktur dan bertahap berdasarkan pada kemiripan sifat antar obyek. Kemiripan sifat tersebut dapat ditentukan dari kedekatan jarak. Secara umum ada dua cara pengelompokan dengan menggunakan metode hirarki, yaitu dengan cara penggabungan (*agglomerative*) dan cara pemisahan (*divisive*) (Everitt dkk, 2011).

Dalam metode hirarki dengan cara penggabungan, pada awal pengelompokan setiap obyek pengamatan dianggap berasal dari kelompok yang berbeda. Kemudian secara bertahap obyek-obyek yang saling berdekatan dikelompokkan. Sehingga pada akhirnya semua obyek berada dalam satu kelompok yang sama (Everitt dkk, 2011).

Sebaliknya adalah metode hirarki dengan cara pemisahan. Langkahnya berlawanan dengan metode sebelumnya. Semua obyek dianggap berasal dari satu kelompok besar. Kemudian dilihat perbedaan antar obyek. Obyek yang berbeda akan dikeluarkan dari kelompok awal dan seterusnya secara bertahap sehingga akan terbentuk kelompok terakhir yang beranggotakan satu obyek saja (Everitt, 2011).

Dari kedua macam metode hirarki tersebut yang paling sering digunakan adalah metode pertama, yaitu metode hirarki dengan cara penggabungan. Algoritma untuk metode ini secara umum adalah sebagai berikut:

1. Membentuk matriks jarak untuk masing-masing obyek
2. Mencari jarak terdekat antara dua obyek. Dua obyek yang mempunyai jarak terdekat digabung di dalam satu kelompok.
3. Menghitung jarak antara obyek-obyek di dalam kelompok yang telah terbentuk dengan obyek di luar kelompok tersebut.
4. Menyusun kembali matriks jarak baru dengan menghilangkan baris dan kolom yang berhubungan dengan anggota kelompok tersebut.
5. Mengulang kembali langkah 2 hingga terbentuk satu kelompok saja.

Menurut Hair dkk (2010), terdapat lima macam algoritma untuk membentuk kelompok dengan metode hirarki, yaitu :

- a. *Single-linkage* (pautan tunggal)

Metode pautan tunggal ini didasarkan atas jarak minimum. Dua obyek yang memiliki jarak terdekat dikelompokkan. Apabila terdapat obyek ketiga yang memiliki jarak terdekat dengan salah satu obyek dalam kelompok yang telah terbentuk, maka obyek tersebut dapat digabung ke dalam kelompok tersebut. Proses ini berlanjut hingga

terbentuk satu kelompok saja. Apabila jarak antara obyek U dengan V dilambangkan dengan  $d_{UV}$ , dan keduanya merupakan anggota dari suatu kelompok, maka jarak kelompok tersebut dengan kelompok atau obyek lain dapat ditentukan sebagai berikut:

$$d_{(UV)W} = \min\{d_{UW}, d_{VW}\} \quad (2.36)$$

b. *Complete linkage* (pautan lengkap)

Metode ini pada dasarnya sama dengan metode pautan tunggal. Hanya saja jarak yang digunakan adalah jarak yang maksimum. Alasan dipilih jarak maksimum adalah agar obyek-obyek yang mempunyai sedikit kesamaan dapat dihubungkan. Jarak antara satu kelompok dengan kelompok atau obyek lain digambarkan sebagai berikut:

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\} \quad (2.37)$$

c. *Average-linkage* (pautan rata-rata)

Metode pautan rata-rata juga memiliki kesamaan dengan dua metode pautan tunggal. Hanya jarak yang digunakan adalah rata-rata jarak dari semua obyek dalam satu kelompok dengan obyek lain di luar kelompok tersebut. Pengelompokan obyek satu dengan yang lain berdasarkan pada rata-rata minimum. Karena menggunakan rata-rata, maka metode ini dianggap lebih stabil, dan tidak bias. Penghitungan jarak antar kelompok adalah dengan menggunakan persamaan:

$$d_{(UV)W} = \text{rata-rata}\{d_{UW}, d_{VW}\} \quad (2.38)$$

dengan  $d_{UW}$  merupakan jarak rata-rata dari U dan W, demikian pula dengan  $d_{VW}$ .

d. *Ward's method*

Dalam metode Ward, perhitungan jarak berdasarkan jumlah kuadrat antara dua kelompok untuk semua peubahnya. Metode ini dapat digunakan apabila jumlah pengamatan tidak terlalu besar. Pada umumnya jarak yang digunakan adalah jarak kuadrat *Euclidean*.

e. *Centroid method* (metode titik pusat)

Jarak yang digunakan dalam metode ini adalah jarak (kuadrat *Euclidean*) antara titik pusat dua kelompok. Titik pusat kelompok ini adalah nilai tengah obyek setiap peubah dalam satu kelompok. Dalam metode ini setiap kali terbentuk kelompok baru, maka titik pusatnya berubah. Keuntungan dari metode ini adalah kecilnya pengaruh pencilan dalam pembentukan kelompok.

Hubungan antar obyek tersebut dapat digambarkan dalam satu buah diagram yang biasa disebut dengan dendrogram, dari dendrogram akan terlihat sebanyak  $N-1$  macam kombinasi kelompok, dengan  $N$

adalah banyaknya obyek yang diamati. Untuk memperoleh banyaknya kelompok yang diinginkan dilakukan pemotongan dendrogram, yaitu pada selisih jarak penggabungan terbesar dimana struktur kelompok selalu stabil untuk jarak yang jauh (Everitt dkk, 2011).

## 2.5. Algoritma *K-Means*

Algoritma *k-means* merupakan salah satu dari metode analisis *cluster* non-hirarki. Algoritma *k-means* mempartisi data yang telah ditentukan oleh peneliti kemudian secara iteratif mengelompokkan objek-objek tersebut kedalam kelompok yang telah ditentukan sampai objek-objek yang dikelompokkan dalam satu *cluster* tidak berpindah posisi lagi. Tidak terdapat pembentukan diagram pada algoritma *k-means*. (Hair dkk, 2010)

Algoritma pengelompokan dengan metode *k-means* adalah sebagai berikut (Hair dkk, 2010):

1. Mempartisi obyek sebanyak  $K$  kelompok (ditentukan oleh peneliti).
2. Menghitung pusat kelompok, dimana pusat kelompok merupakan rata-rata dari keseluruhan objek yang berada dalam kelompok tersebut.
3. Menghitung jarak setiap obyek ke pusat kelompok.
4. Menentukan obyek yang memiliki jarak terdekat dengan pusat kelompok. Jika obyek berpindah dari posisi awal (langkah a) maka pusat kelompok harus dihitung kembali.
5. Mengulangi langkah 2-5 sampai tidak ada lagi obyek yang berpindah posisi.

## 2.6. Pemilihan Metode Terbaik

Analisis *cluster* merupakan analisis statistika yang bertujuan mengelompokkan kumpulan objek ke dalam sejumlah kelompok. Pengelompokan bertujuan mengelompokkan objek-objek yang memiliki kemiripan karakteristik ke dalam satu kelompok (Everitt dkk, 2011). Objek-objek dalam kelompok memiliki kemiripan yang dekat, sehingga semakin mirip karakteristik objek dalam kelompok-kelompok maka pengelompokan dilakukan semakin tepat. Ketepatan pengelompokan dapat diukur dengan beberapa cara diantaranya dengan menghitung kepadatan kelompok dan koefisien korelasi *cophenetic*:

### 2.6.1. Korelasi *Cophenetic*

Dendrogram merupakan representasi secara matematis dan visual dari prosedur pengelompokan secara lengkap. Dendrogram dihasilkan jika pengelompokan dilakukan dengan menggunakan analisis *cluster* hirarki. Bentuk dendrogram identik dengan diagram pohon, titik titik pada dendrogram menunjukkan *cluster* sedangkan panjang batang (*height*) merepresentasikan jarak dimana objek-objek *digabungkan* dalam kelompok (Everitt dkk, 2011).

Dendrogram sebagai representasi proses pengelompokan analisis *cluster* secara hirarki akan membentuk struktur yang berbeda tergantung pada jarak maupun *linkage* yang dipakai pada proses pengelompokan. Struktur hirarki data yang dihasilkan perlu diidentifikasi ketepatannya, apakah struktur tersebut telah sesuai dengan hubungan nyata antar objek yang dapat dilihat melalui kedekatan observasi masing-masing objek. Kesesuaian struktur data yang dihasilkan oleh analisis *cluster* hirarki dengan kedekatan observasi antar objek dapat diketahui dengan menghitung koefisien korelasi *cophenetic* pengelompokan tersebut (Everitt dkk, 2011).

Korelasi *cophenetic* adalah korelasi antara matriks *cophenetic* dengan matriks jarak antar objek. Matriks *cophenetic* adalah matriks yang elemen-elemennya merupakan jarak dimana objek-objek digabungkan ke dalam satu kelompok (*height*). Jika komponen matriks dituliskan sebagai  $D_{ij}$ , maka  $D_{ij}$  adalah jarak dimana dua objek bergabung pada *cluster* yang sama pada dendrogram (Everitt dkk, 2011). Sebagai contoh jika diketahui sebuah matriks jarak

$$D_{ij} = \begin{matrix} 1 & \left( \begin{array}{cccccc} 0.0 & & & & & \\ 2.0 & 0.0 & & & & \\ 6.0 & 5.0 & 0.0 & & & \\ 10.0 & 9.0 & 4.0 & 0.0 & & \\ 9.0 & 8.0 & 5.0 & 3.0 & 0.0 & \end{array} \right) \end{matrix} \quad (2.39)$$

Matriks jarak pada persamaan 2.39 merupakan jarak dari objek satu sampai objek lima, objek-objek tersebut dikelompokkan menggunakan analisis *cluster* hirarki dengan *single linkage*. Sehingga didapatkan jarak pengelompokannya pada Tabel 2.3.

Tabel 2.1 Contoh jarak penggabungan tiap stage

Stage	Jarak penggabungan	Objek yang digabungkan
1	2,0	1 dan 2
2	3,0	4 dan 5
3	4,0	(4,5) dan 3
4	5,0	(1,2) dan (3,4,5)

Dari Tabel 2.1 dapat dibuat matriks *cophenetic* dari pengelompokan lima objek dari data contoh sebagai berikut :

$$C_{ij} = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{pmatrix} 0.0 & & & & & \\ 2.0 & 0.0 & & & & \\ 5.0 & 5.0 & 0.0 & & & \\ 4.0 & 5.0 & 5.0 & 4.0 & 0.0 & \\ 5.0 & 5.0 & 4.0 & 3.0 & 0.0 & \end{pmatrix} \quad (2.40)$$

Korelasi *cophenetic* merupakan korelasi antar elemen pada matriks diagonal jarak dengan matriks diagonal *cophenetic*. Koefisien korelasi *cophenetic* dari contoh pengelompokan dihitung menggunakan rumus:

$$r_{XY} = \frac{N \cdot \sum X \cdot Y - (\sum X)(\sum Y)}{\sqrt{[N \cdot \sum X^2 - (\sum X)^2][N \cdot \sum Y^2 - (\sum Y)^2]}} \quad (2.41)$$

dengan :

- $r_{XY}$  = Koefisien korelasi antara matriks  $D_{ij}$  dan matriks  $C_{ij}$
- $\sum XY$  = Jumlah perkalian antara elemen matriks  $D_{ij}$  dan matriks  $C_{ij}$
- $\sum X^2$  = Jumlah dari kuadrat nilai elemen matriks  $D_{ij}$
- $\sum Y^2$  = Jumlah dari kuadrat nilai elemen matriks  $C_{ij}$
- $(\sum X)^2$  = Jumlah elemen matriks  $D_{ij}$  kemudian dikuadratkan
- $(\sum Y)^2$  = Jumlah nilai elemen matriks  $C_{ij}$  kemudian dikuadratkan

### 2.6.2. Silhouette Width (Koefisien Silhouette)

Kauffman dan Rousseuw (1990) memperkenalkan suatu metode untuk melihat kepadatan dan kepadatan *cluster* secara grafis melalui siluet (*silhouettes*). Siluet adalah gambaran, atau karakteristik dari suatu objek atau benda yang digambarkan direpresentasikan melalui bentuk yang solid dari satu warna. Setiap *cluster* diwakili oleh satu siluet, yang

menunjukkan objek-objek yang berada dalam satu *cluster* sudah dikelompokkan baik atau tidak. Seluruh pengelompokan ditunjukkan dengan memplot semua siluet menjadi diagram tunggal, yang memungkinkan untuk membandingkan kualitas cluster. Metode siluet ini dapat dilakukan jika data yang dikelompokkan berupa data rasio dengan batas  $k \neq 1$  dan  $k \neq n$ .

Siluet dibentuk dengan memplotkan  $s(i)$  dari masing masing objek  $i$ .  $s(i)$  dapat dihitung melalui rumus (Kauffman dan Rousseeuw, 1990) :

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (2.42)$$

$a(i)$  didefinisikan sebagai rata-rata jarak dari objek  $i$  terhadap semua anggota *cluster* A, dimana *cluster* A beranggotakan selain objek  $i$  dan tidak beranggotakan anggota tunggal. Sedangkan  $b(i)$  adalah jarak terkecil dari semua objek  $i$  terhadap semua anggota *cluster* B, dimana *cluster* merupakan *nearest neighbor* dari objek  $i$ , dengan kata lain jika *cluster* A dihilangkan maka *cluster* B menjadi *cluster* terdekat dengan objek  $i$ , dapat dikatakan jika *cluster* B merupakan pilihan kedua terbaik untuk objek  $i$  (Kauffman dan Rousseeuw, 1990), dimana  $s(i)$  berada pada rentang :

$$-1 \leq s(i) \leq 1 \quad (2.43)$$

Nilai  $s(i)$  mendekati 1 menunjukkan bahwa jarak antar anggota *cluster* ( $a(i)$ ) jauh lebih kecil dibandingkan jarak antar objek dalam *cluster* dengan *cluster* lainnya, sehingga dikatakan pengelompokan dilakukan dengan baik. Sebaliknya jika nilai  $s(i)$  mendekati -1 menunjukkan bahwa jarak antar anggota *cluster* ( $a(i)$ ) jauh lebih besar dibandingkan jarak antar objek dalam *cluster* dengan *cluster* lainnya, sehingga dikatakan bahwa terjadi *misclassified* (salah pengelompokan). Nilai  $s(i)$  sama dengan nol menunjukkan bahwa terdapat keragu-raguan dalam pengelompokan yang dilakukan (Kauffman dan Rousseeuw, 1990).

Kauffman dan Rousseeuw (1990) membuat sebuah acuan dalam menyatakan kualitas pengelompokan yaitu dengan menggunakan koefisien *silhouette* yang diperoleh melalui persamaan :

$$SC = \max_k \bar{s}(k) \quad (2.44)$$

$\bar{s}(k)$  adalah rata-rata nilai siluet tiap *cluster* yang terbentuk dalam pengelompokan. Kriteria Subjektif kualitas pengelompokan berdasarkan nilai koefisien *silhouette* terdapat pada Tabel 2.2

Tabel 2.2 Kriteria Subjektif Kualitas Pengelompokk berdasarkan Koefisien *Silhouette* (SC)

Nilai SC	Interpretasi Oleh Kauffman
0,71 - 1,00	<i>Strong Classification</i>
0,51 - 0,70	<i>Good Classification</i>
0,26 – 0,50	<i>Weak Classification</i>
0 - 0,25	<i>Bad Classification</i>

Kriteria yang ditetapkan oleh Kauffman dan Rousseuw (1990) bukan patokan resmi untuk menentukan kualitas *cluster*, perbedaan kondisi dan persepsi peneliti dapat memunculkan kriteria yang berbeda (Everitt, 2011)

## 2.7. Tinjauan Non-Statistika

### 2.7.1. Pengertian, Fungsi, dan Tujuan Bank

Pengertian bank disebutkan dalam Undang-Undang No.10 Tahun 1998 Pasal 1 huruf 2 yaitu “ Bank adalah badan usaha yang menghimpun dana dari dalam masyarakat dalam bentuk simpanan dan menyalurkannya kepada masyarakat dalam bentuk kredit dan atau bentuk bentuk lainnya dalam rangka meningkatkan taraf hidup rakyat banyak”(Suhardi, 2013).

Dari defini tentang bank yang telah disebutkan dalam undang-undang dapat diketahui bahwa bank berfungsi sebagai “*financial intermediary*” yaitu, badan usaha yang yang tugas utamanya melakukan penghimpunann dana dari masyrakat dan kemudian menyalurkannya kembali kepada masyarakat serta memberikan jasa-jasa lainnya dalam lalu lintas pembayaran. Bank selain sebagai badan usaha juga sekaligus sebagai lembaga keuangan, yang juga mempunyai kewajiban untuk menjaga kestabilan nilai uang, mendorong kegiatan ekonomi, dan perluasan kesempatan kerja (Usman, 2003).

Fungsi dan tujuan perbankan nasional dalam kehidupan nasional bangsa Indonesia adalah sebagai berikut (Usman, 2003):

1. Fungsi bank adalah sebagai “*financial intermedairy*” yang mempunyai usaha utama dalam melakukan penghimpunan dan penyaluran dana dari masyarakat atau pemidahan dana masyarakat dari unit surplus kepana unit defisit atau dengan kata lain sebagai sarana pemindahan uang dari penabung kepada peminjam

2. Penghimpunan dan penyaluran dana masyarakat bertujuan untuk menunjang tugas penyelenggaraan negara, antara lain :
  - a. Melaksanakan pembangunan nasional dengan kata lain sebagai *agen of development*
  - b. Mewujudkan trilogi pembangunan, antara lain:
    - 1) Meningkatkan pemerataan kesejahteraan seluruh rakyat
    - 2) Meningkatkan pertumbuhan ekonomi nasional
    - 3) Meningkatkan stabilitas nasional yang sehat dan dinamis serta meningkatkan taraf hidup dan kesejahteraan seluruh rakyat Indonesia
3. Menerapkan prinsip kehati-hatian (*prudential banking*) untuk melindungi secara baik apa yang dititipkan masyarakat kepada bank tersebut dengan cara :
  - a. Efisien, sehat , wajar dalam persaingan yang sehat yang semakin mengglobal.
  - b. Menyalurkan dana masyarakat tersebut ke bidang-bidang yang produktif, bukan konsumtif
4. Memenuhi ketentuan dan syarat sebagai bank yang sehat, hal ini berfungsi untuk mencegah terjadinya praktek-praktek yang merugikan masyarakat.

### **2.7.2. Jenis Bank**

Berdasarkan pasal 5 Undang-Undang No 7 Tahun 1992, kelembagaan bank dibagi menjadi dua jenis bank, yaitu bank Umum dan Bank Perkreditan Rakyat (BPR) (Usman, 2003):

#### **1. Bank Umum**

Definisi Bank Umum menurut Undang-Undang No. 10 Tahun 1998 adalah bank yang melaksanakan kegiatan usaha secara konvensional dan atau berdasarkan prinsip syariah yang secara kegiatannya memberikan jasa dalam lalu lintas pembayaran (*commercial bank*).

Bank umum mempunyai tugas pokok dalam menghimpun dana dari masyarakat, memberikan pinjaman kepada masyarakat, dan memberikan jasa melalui mekanisme keuangan kepada masyarakat. Kegiatan yang dilakukan Bank Umum anatara lain adalah (Ahman dan Indriani, 2007):

- a. Memberi dan menerima pinjaman dari perusahaan lain atau masyarakat.

- b. Menerima titipan barang-barang berharga.
  - c. Melakukan kegiatan dalam valuta asing.
  - d. Melayani jasa pengiriman (*transfer*) antar bank.
  - e. Melakukan giro dan inkaso antar bank, inkaso adalah melaksanakan amanat dari pihak ketiga dengan melakukan penagihan sejumlah uang kepada seseorang atau badan tertentu di kota lain yang telah ditunjuk oleh si pemberi amanat
2. Bank Perkreditan Rakyat (BPR)

Bank Perkreditan Rakyat (BPR) adalah bank yang melaksanakan kegiatan usaha secara konvensional atau berdasarkan prinsip syariah yang dalam kegiatannya tidak memberikan jasa dalam lalu lintas pembayaran. BPR tidak ikut serta dalam memberikan jasa lalu lintas pembayaran. BPR hanya menerima bentuk tabungan dan deposito (Ahman dan Indriani, 2007).

### 2.7.3. Sumber Dana Bank

Sumber dan ayang dimiliki oleh bank dibedakan menjadi tiga jenis yaitu (Dendawijaya, 2005) :

1. Dana Pihak Pertama (dana sendiri), adalah dana yang berasal dari pemegang saham atau milik bank.
2. Dana Pihak Kedua (dana pinjaman), adalah dana yang berasal dari pihak luar.
3. Dana Pihak Ketiga (dana masyarakat), adalah dana yang berasal dari masyarakat (perorangan maupun badan usaha). Dana ini dihimpun dalam bentuk giro, deposito dan tabungan.

a. Giro (*Demand Deposits*)

Menurut Undang-Undang No. 10 Tahun 1998, giro adalah simpanan atau dana pihak ketiga, dimana penarikannya dapat dilakukan setiap saat dengan media yaitu cek, bilyet giro dan sarana perintah pembayaran lainnya.

b. Deposito (*Time Deposits*)

Menurut Undang-Undang No. 10 Tahun 1998 pasal 1 ayat 7, deposito adalah simpanan yang penarikannya hanya dalam jangka waktu tertentu berdasarkan perjanjian nasabah yang telah dilakukan sebelumnya.

c. Tabungan (*Savings*)

Menurut Undang-Undang No.10 Tahun 1998, tabungan adalah simpanan pihak ketiga yang penarikannya hanya dapat

dilakukan menurut syarat-syarat tertentu yang telah disepakati, tetapi tidak dapat ditarik dengan cek, bilyet giro dan atau yang lainnya yang sejenis dengan itu.

Menurut Dendawijaya (2005) Dana Pihak Ketiga (DPK) merupakan sumber dana terbesar yang dimiliki oleh bank. DPK ini bisa mencapai 80%-90% dari seluruh dana yang dikelola bank. Kegiatan bank setelah melakukan penghimpunan dana dari masyarakat dalam bentuk tabungan, giro, deposito adalah menyalurkan kembali dalam bentuk kredit. Bank menghasilkan keuntungan dari aktivitasnya menyalurkan kredit. Data statistik Perbankan Nasional Indonesia mencatat jumlah Dana pihak Ketiga (DPK) perbankan nasional, Bank Umum menghasilkan jumlah DPK terbesar dibandingkn Bank Syariah dan Bank Perkreditan Rakyat (BPR). Sehingga bank umum memiliki andil besar dalam menggerakkan roda perekonomian Indonesia. Komposisi data DPK Perbankan Nasional dari tahun 2004 hingga 2013 pada Bulan Desember adalah sebagai berikut :

Tabel 2.3 Komposisi dana Pihak Ketiga (DPK) Perbankan Nasional Periode 2004-2013 (Bulan Desember) dalam Milyar Rupiah

Tahun	Bank Umum	Bank Syariah	BPR
2004	963106	11718	11161
2005	1127937	15581	13178
2006	1287102	19347	15771
2007	1510834	28011	18719
2008	1753292	36852	21339
2009	1950712	52271	25552
2010	2338824	76036	31312
2011	2785024	115415	38209
2012	3225198	147512	44870
2013	3663938	147512	50520
Total	16942029	650255	270631
Proporsi	94,8%	3,64%	1,51%

Sumber : Data Statistik Perbankan Indonesia (Data diolah)

Berdasarkan data Statistik Perbankan Indonesia. Rata-rata presentasi DPK Bank Umum terhadap Perbankan Nasional pada tahun 2004 hingga tahun 2013 (posisi Desember) adalah 94,8%, hal ini menunjukkan bahwa jumlah DPK Bank Umum jauh di atas Bank Syariah dan BPR.

## BAB III METODE PENELITIAN

### 3.1. Sumber Data

Data yang digunakan pada penelitian ini adalah data sekunder yaitu jumlah Dana Pihak Ketiga seluruh provinsi di Indonesia Bulan Januari 2005 – Desember 2013 kecuali Provinsi Irian Jaya Barat, Kepulauan Riau, dan Sulawesi Barat karena baru dibentuk sehingga tidak memadai untuk pengambilan data. Data diambil dari Bank Indonesia ([bi.go.id](http://bi.go.id)). Data pengamatan yang akan digunakan adalah pada 30 Provinsi di Indonesia. Data Dana Pihak Ketiga 30 Provinsi di Indonesia dilihat pada Lampiran 1.

### 3.2. Langkah-langkah Penelitian

Dalam penulisan skripsi ini, peneliti menggunakan beberapa software yaitu *Minitab*, *R console* dan *Microsoft Excell*. Berikut Proses analisis *cluster* dalam penulisan skripsi ini:

1. Melakukan pengambilan data, dimana data yang diambil sudah cukup mewakili observasi.
2. Uji stasioneritas terhadap ragam dilakukan sebagai asumsi pengelompokan data deret waktu.
3. Apabila data deret waktu tidak stasioner maka dilakukan transformasi terlebih dahulu agar deret waktu menjadi stasioner, jika sudah stasioner maka langsung dilakukan uji stasioneritas terhadap rata-rata.
4. Uji stasioneritas terhadap rata-rata dilakukan sebagai asumsi pengelompokan data deret waktu.
5. Apabila data deret waktu tidak stasioner maka dilakukan differensiasi hingga data stasioner terhadap rata-rata, jika sudah stasioner maka langsung dapat dilakukan perhitungan jarak.
6. Melakukan perhitungan jarak yaitu jarak *Autocorrelation-based*, *Dynamic Time Warping*, *Periodogram-based*, dan *Complexity invariant*
7. Melakukan analisis kelompok hirarki (*average linkage*) sesuai matriks jarak yang telah didapat.
8. Kemudian dilakukan pemilihan metode terbaik dengan menggunakan korelasi *cophenetic*.

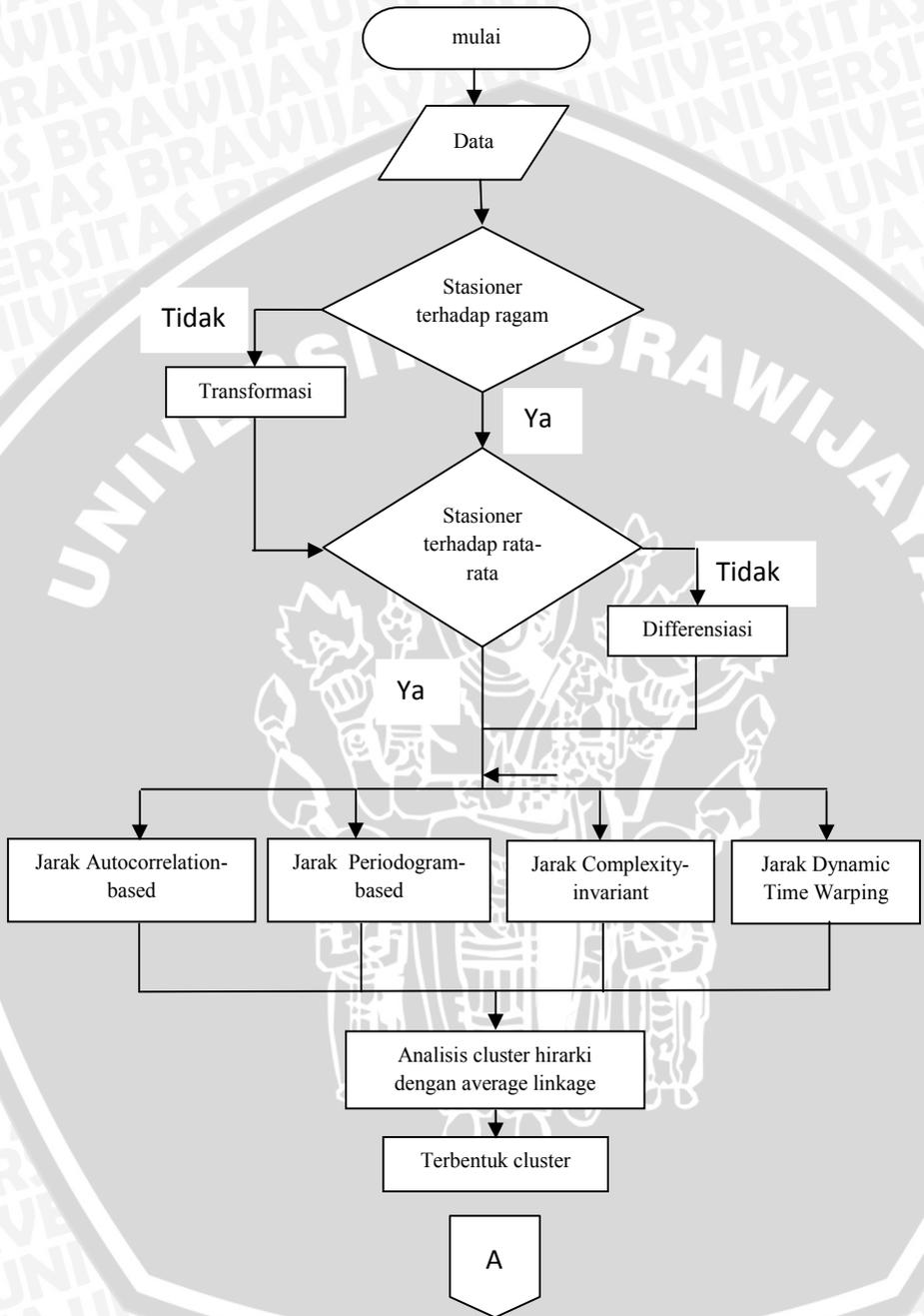
9. Melakukan analisis *cluster* menggunakan algoritma *k-means* dengan jumlah dan pusat *cluster* yang didapat melalui analisis *cluster* hirarki.
10. Melakukan perhitungan koefisien *silhouette* untuk melihat apakah jumlah kelompok dan anggota kelompok sudah tepat atau tidak.
11. Kemudian interpretasi dan selesai.

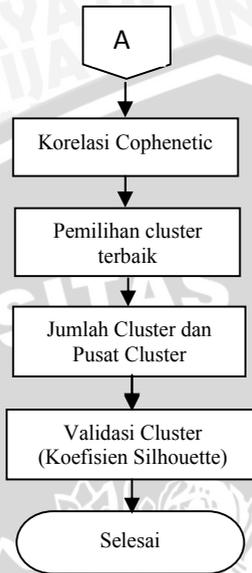
Langkah-langkah analisis *cluster* data deret waktu dan validasi *cluster* disajikan pada Gambar 3.1

Adapun langkah-langkah untuk menentukan *optimal warping path* yang digunakan untuk jarak *Dynamic Time Warping* dari dua deret waktu adalah sebagai berikut :

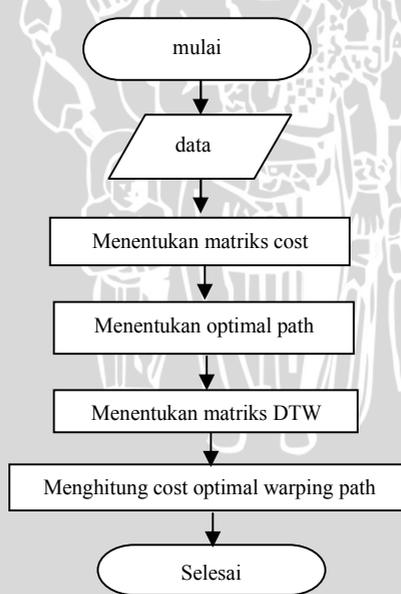
1. Menentukan matriks *cost*, bentuk matriks *cost* sesuai dengan persamaan 2.30, sedangkan untuk perhitungan masing masing komponen matriksnya menggunakan persamaan 2.31
2. Menentukan *optimal path*
3. Menentukan matriks DTW dengan masing masing komponennya dapat dihitung menggunakan persamaan 2.32, 2.33, 2.34
4. Menghitung *cost optimal warping path* menggunakan persamaan 2.35, yang akan digunakan sebagai jarak DTW dari dua deret waktu

Langkah-langkah perhitungan jarak *Dynamic Time Warping* (DTW) antara dua deret waktu disajikan pada Gambar 3.2





**Gambar 3.1** Diagram Alir Analisis Kelompok dengan jarak *autocorrelation-based, periodogram-based, CID, DTW*



**Gambar 3.2** Diagram alir perhitungan jarak *Dynamic Time Warping*

## BAB IV PEMBAHASAN

### 4.1 Deskripsi Data

Data Dana Pihak Ketiga (DPK) yang digunakan pada penelitian ini adalah data DPK dari 30 provinsi di Indonesia pada bulan Januari 2005 – Desember 2013. Statistik deskriptif pada data DPK 30 provinsi dapat dilihat pada Tabel 4.1.

Tabel 4.1 Statistik Deskriptif DPK 30 Provinsi di Indonesia Bulan Januari 2005 – Desember 2013 dalam Milyar Rupiah

Provinsi	Rata-rata	Minimum	Maksimum
Jawa Barat	158338	74731	300571
Banten	46391	18877	105532
DKI Jakarta	989080	495897	1860446
DI Yogyakarta	20198	10199	37073
Jawa Tengah	92398	45367	167398
Jawa Timur	187412	93969	335308
Bengkulu	4509	1649	8379
Jambi	11527	5063	19521
Aceh	17938	7702	26141
Sumatra Utara	91148	43821	155204
Sumatra Barat	16830	7731	27070
Riau	36299	22649	56878
Sumatra Selatan	33427	13951	54244
Bangka Belitung	7646	2838	12774
Lampung	15739	6209	26116
Kalimantan Selatan	18909	6688	36447
Kalimantan Barat	19922	8566	36273
Kalimantan Timur	46310	16992	88869
Kalimantan Tengah	8805	3248	16738
Sulawesi Tengah	6445	2702	11364
Sulawesi Selatan	32467	14142	61163
Sulawesi Utara	9903	4196	17720
Gorontalo	1999,9	732	3361
Sulawesi Tenggara	5696	1830	11665
Nusa Tenggara Barat	7526	3131	14648

Tabel 4.1 Lanjutan

Provinsi	Rata-rata	Minimum	Maksimum
Bali	32942	15159	64234
Nusa Tenggara Timur	9363	4130	16694
Maluku	5276	2559	9379
Papua	17077	6083	30255
Maluku utara	2803	688	5012

Berdasarkan Tabel 4.1 dapat diketahui selama Bulan Januari 2005 sampai dengan bulan Desember 2013 jumlah DPK terbesar terdapat pada Provinsi DKI Jakarta dengan kisaran 989080 milyar perbulannya. Hal ini dikarenakan jumlah bank umum di Provinsi DKI Jakarta lebih banyak dari pada provinsi lainnya, begitu juga dengan jumlah nasabahnya. Jumlah minimum Dana Pihak Ketiga di DKI Jakarta yaitu terdapat pada bulan Januari 2005 sebesar 495897 milyar dan mencapai jumlah maksimum pada bulan Desember 2013 yaitu sebesar 1860446 milyar.

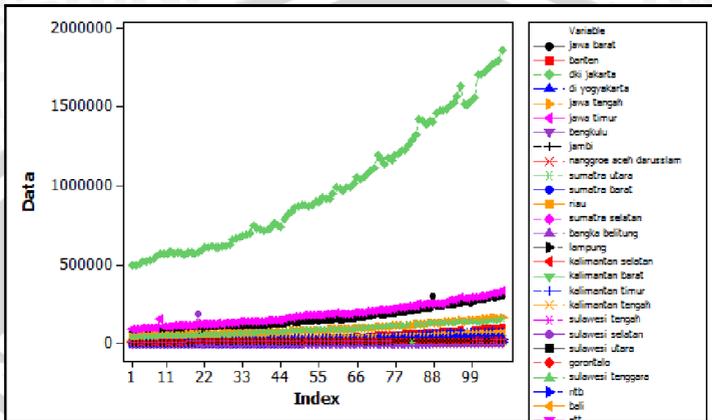
Dana Pihak Ketiga (DPK) terkecil selama Januari 2005 sampai dengan Desember 2013 terdapat pada Provinsi Maluku Utara dengan kisaran 2803 milyar. Hal ini dikarenakan minimnya jumlah cabang bank umum di Maluku Utara dibandingkan provinsi lainnya. Jumlah minimum Dana Pihak Ketiga pada Provinsi Maluku Utara terdapat pada bulan Agustus 2005 yaitu sebesar 688 milyar dan mencapai jumlah maksimum sebesar 5012 pada bulan November 2013.

Berdasarkan Tabel 4.1 dan plot DPK tiap provinsi pada Lampiran 1 dapat dikatakan bahwa setiap tahunnya jumlah DPK meningkat, hal ini dapat dilihat dari jumlah minimum DPK dari tiap provinsi yang selalu berada pada tahun 2005 dan jumlah maksimum hampir selalu berada pada tahun 2013. Hal ini dikarenakan setiap tahunnya di tiap provinsi terdapat penambahan jumlah nasabah bank umum, dan kenaikan nilai mata uang di tiap tahunnya sehingga jumlah Dana Pihak Ketiga dari tiap provinsi mengalami peningkatan setiap tahunnya.

#### 4.2 Plot Deret Waktu

Langkah awal dalam melihat pola data deret waktu adalah dengan membuat plot data deret waktu untuk mengetahui tipe dan karakteristik data deret waktu tersebut secara visual. Plot data

deret waktu dapat dilihat pada Gambar 4.1, sedangkan plot data deret waktu tiap provinsi dapat dilihat dalam Lampiran 2.



Gambar 4.1 Plot DPK 30 Provinsi di Indonesia

Dari Gambar 4.1 dapat dilihat bahwa semua provinsi mempunyai pola data yang hampir sama yaitu membentuk tren positif. DKI Jakarta mempunyai Jumlah DPK terbesar dari 29 provinsi lainnya berada diatas plot deret waktu provinsi lainnya. Terdapat peningkatan jumlah DPK selama selang waktu bulan Januari 2005 – Desember 2013. Karakteristik dari data tren adalah ketidakstasioneritasan terhadap rata-rata, dan seringkali juga tidak stasioner terhadap ragam. Untuk melihat lebih jelas apakah data deret waktu dari setiap provinsi stasioner atau tidak dapat dilakukan dengan menggunakan uji *Augmented Dickey Fuller* dan transformasi *Box-Cox*.

### 4.3 Pengujian Stasioneritas Data DPK 30 Provinsi di Indonesia

Pengujian stasioneritas pada data deret waktu dilakukan sebanyak dua kali. Pengujian yang pertama untuk mengetahui apakah data deret waktu tersebut telah stasioner terhadap ragam atau belum, yang kedua untuk mengetahui apakah data deret waktu tersebut telah stasioner terhadap rata-rata atau belum. Apabila data sudah stasioner terhadap ragam dan rata-rata maka data tersebut dapat dikelompokkan menggunakan Analisis *Cluster* data deret waktu.

### 4.3.1 Pengujian Stasioneritas Terhadap Ragam

Pengujian stasioneritas dapat dilakukan dengan melihat nilai lambda ( $\lambda$ ) apabila nilai lambda ( $\lambda$ ) = 1 maka data deret waktu sudah stasioner terhadap ragam, namun apabila nilai lambda ( $\lambda$ ) < 1 maka dilakukan transformasi *Box-Cox* hingga data deret waktu tersebut stasioner terhadap ragam. Nilai lambda untuk tiap provinsi dapat dilihat pada Lampiran 3. Data DPK 30 provinsi di Indonesia tidak stasioner terhadap ragam, sehingga harus dilakukan transformasi untuk menstasionerkan data deret waktu tersebut.

Nilai lambda ( $\lambda$ ) dari 30 provinsi < 1 sehingga data dari 30 provinsi dikatakan belum stasioner terhadap ragam dengan tingkat kepercayaan 95%. Sehingga perlu dilakukan transformasi agar data deret waktu menjadi stasioner terhadap ragam

### 4.3.2 Pengujian Stasioneritas Terhadap Rata-rata

Pengujian stasioneritas dapat dilakukan dengan Uji *Augmented Dickey Fuller*. Kriteria pengujian ini adalah apabila nilai  $|ADF \text{-test}| > |\text{titik kritis}|$  maka data tidak mengandung unit root atau dapat dikatakan data telah stasioner. Jika data tidak stasioner maka dapat dilakukan *differencing*. Hasil pengujian *Augmented Dickey Fuller* ini dapat dilihat pada Lampiran 4. Lampiran 4 menunjukkan bahwa data Dana Pihak Ketiga dari 30 provinsi belum stasioner terhadap rata-rata

Pada Lampiran 4 menunjukkan bahwa nilai *p-value* masing-masing provinsi pada masing-masing daerah lebih dari  $\alpha$  (0,05) sehingga 30 provinsi tersebut dinyatakan tidak stasioner rata-rata. *Differencing* perlu dilakukan pada data yang tidak stasioner terhadap rata-rata hingga data tersebut stasioner. *Differencing* dilakukan sampai nilai *p-value* dari masing-masing provinsi lebih kecil dari  $\alpha$  (0,05).

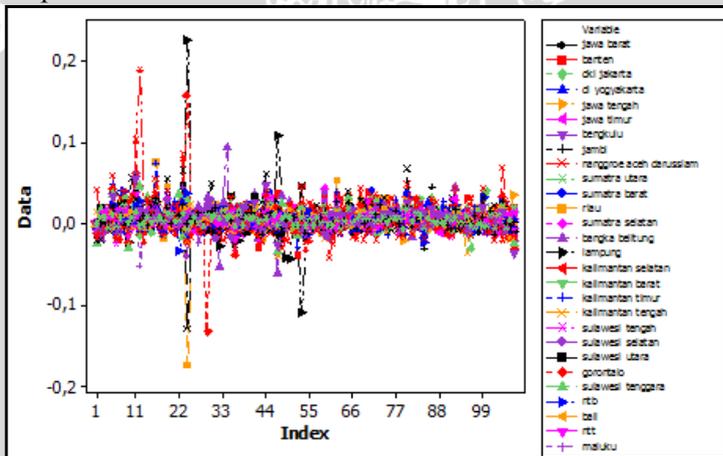
## 4.4 Stasioneritas Data Deret Waktu

Stasioneritas merupakan syarat penting dalam melakukan analisis *cluster* pada data deret waktu, sehingga apabila data deret waktu belum stasioner maka perlu distasionerkan terlebih dahulu. Apabila data tidak stasioner terhadap ragam maka dilakukan transformasi *Box-Cox* hingga nilai lambda ( $\lambda$ ) = 1 (stasioner), sedangkan apabila data tidak stasioner terhadap rata-rata maka dilakukan *differencing* hingga nilai *p-value* pada uji *Augmented Dickey Fuller* lebih kecil dari  $\alpha$  (0,05). Apabila sejumlah data deret waktu tidak stasioner terhadap ragam dan

rata-rata, maka transformasi *box-cox* dan *differencing* bisa dilakukan secara bersamaan pada data deret waktu.

Pengujian *Augmented Dickey Fuller* dilakukan kembali pada data hasil transformasi masing-masing provinsi. Hasil uji *Augmented Dickey Fuller* dapat dilihat pada Lampiran 6. Pada Lampiran 6 dapat dilihat bahwa data deret waktu dari semua provinsi telah stasioner.

Data Dana Pihak Ketiga (DPK) dari 30 provinsi yang telah ditransformasi menunjukkan nilai *p-value* yang lebih kecil dari  $\alpha = 0,05$ . Hal ini menunjukkan bahwa data telah stasioner terhadap rata-rata, dan ragam pada data deret waktu tersebut telah dianggap stabil setelah dilakukan transformasi log pada data hasil *differencing* (Suhartono, 2008). Plot data deret waktu yang telah ditransformasi dapat dilihat pada Gambar 4.2. Pada Gambar 4.2 dapat dilihat plot deret waktu tiap provinsi menyebar secara merata tidak membentuk pola melebar atau menyempit.



Gambar 4.2 Plot Data Dana Pihak Ketiga setelah ditransformasi

Menurut Gambar 4.2 data hasil transformasi sudah menunjukkan ciri-ciri data yang stasioner yaitu tidak membentuk tren naik atau turun dan tidak menunjukkan pola gelombang, melebar maupun menyempit. Plot data deret waktu hasil transformasi masing-masing provinsi dapat dilihat pada Lampiran 5. Plot deret waktu sudah datar dan membentuk gelombang, maka data deret waktu telah stasioner terhadap ragam.

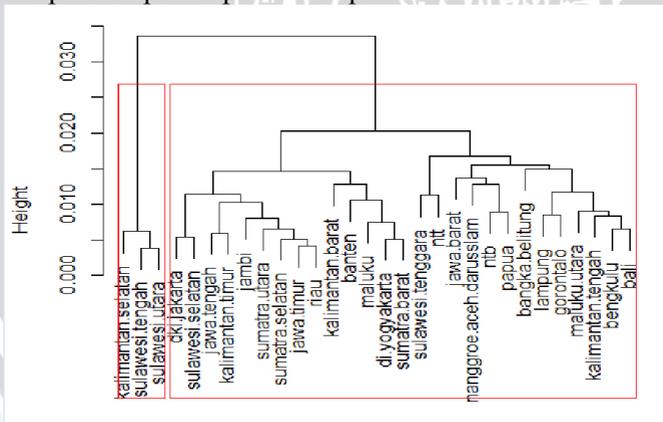
## 4.5 Analisis Cluster Deret Waktu

Data deret waktu yang telah stasioner terhadap ragam dan rata-rata dapat dikelompokkan dengan menggunakan jarak-jarak tertentu atau melalui pendekatan tertentu, dalam penelitian ini data deret waktu akan dikelompokkan dengan menggunakan 4 jarak, yaitu jarak *autocorrelation-based*, *periodogram-based*, *complexity invariant*, dan *dynamic time warping*.

### 4.5.1 *Autocorrelation-based Distance*

Jarak *Autocorrelation-based* merupakan jarak yang didapat melalui fungsi autokorelasi yang dihasilkan. Masing-masing nilai fungsi autokorelasi yang didapat dari data deret waktu dihitung jaraknya menggunakan jarak *euclidean* sesuai dengan persamaan 2.28. Setelah didapatkan matriks jarak maka dilakukan pengelompokan secara hirarki. Jumlah kelompok yang dihasilkan pada analisis *cluster* hirarki didapatkan dengan menghitung selisih jarak penggabungan terbesar.. Jarak penggabungan tiap *stage* (tahapan) dan selisih antar tiap jarak penggabungannya dapat dilihat pada Lampiran 7.

Selisih terbesar pada jarak penggabungan antara *stage* 29 dengan *stage* 28 yaitu sebesar 0,01338366 sehingga ditetapkan jumlah kelompok yang terbentuk sebanyak  $30 - 28 = 2$  kelompok. Anggota masing masing kelompok dapat diketahui melalui pemotongan dendrogram, dengan memotong dendrogram menjadi dua kelompok. Anggota tiap kelompok dapat dilihat pada Gambar 4.3.



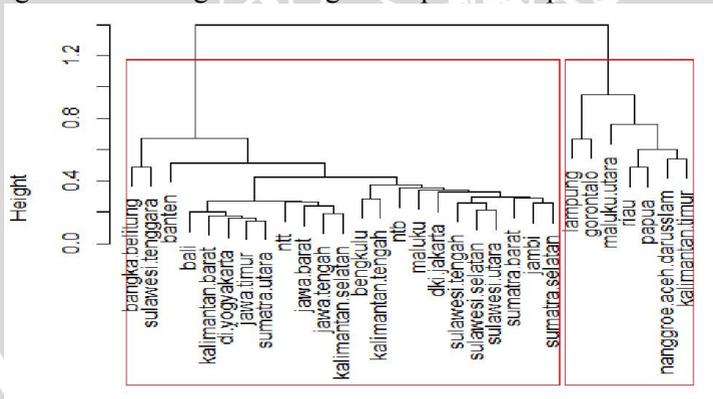
Gambar 4.3 Dendrogram Analisis Cluster Hirarki Menggunakan *Auocorrelation-based Distance*

Berdasarkan Gambar 4.3 dapat diketahui bahwa dengan menggunakan *autocorrelation-based* pada analisis *cluster* hirarki, didapatkan dua kelompok deret waktu. Adapun anggota *cluster* pertama antara lain : Kalimantan Selatan, Sulawesi Selatan, Sulawesi Utara, sedangkan anggota *cluster* kedua antara lain Bangka Belitung, Sulawesi Tenggara, Banten, Bali, Kalimantan Barat, DI Yogyakarta, Jawa Timur, Sumatra Utara, NTT, Jawa Barat, Jawa Tengah, Bengkulu, Kalimantan Tengah, NTB, Maluku, DKI Jakarta, Sulawesi Tengah, Sumatra Barat, Jambi, Sumatra Selatan, Lampung, Gorontalo, Maluku Utara, Riau, Papua, Nanggroe Aceh Darussalam, dan Kalimantan Timur.

#### 4.5.2 Complexity Invariant Distance

Jarak *complexity invariant* merupakan Jarak yang menggunakan informasi tentang perbedaan kompleksitas antara dua deret waktu sebagai faktor koreksi dari jarak yang telah ada, dalam kasus ini yang digunakan merupakan jarak *euclidean*, setelah dilakukan perhitungan jarak dan kemudian dilakukan pengelompokan secara hirarki. Jumlah kelompok ditentukan melalui pemotongan dendrogram pada selisih terbesar antara *stage* penggabungan yang dapat dilihat pada Lampiran 8.

Selisih terbesar pada jarak penggabungan anatara stage 29 dengan stage 28 yaitu sebesar 0,447124 sehingga ditetapkan jumlah kelompok yang terbentuk sebanyak  $30 - 28 = 2$  kelompok. Anggota masing masing kelompok dapat diketahui melalui pemotongan dendrogram. Pemotongan dendrogram dapat dilihat pada Gambar 4.4.



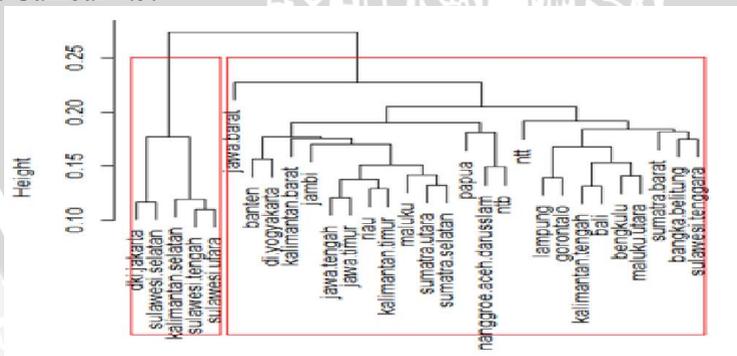
Gambar 4.4 Dendrogram Analisis *Cluster* Hirarki Menggunakan *Complexity Invariant Distance*

Berdasarkan Gambar 4.4 dapat diketahui bahwa dengan menggunakan *Complexity Invariant Distance* (CID) pada analisis *cluster* hirarki, didapatkan dua kelompok deret waktu. Adapun anggota *cluster* pertama antara lain : Bangka Belitung, Sulawesi Tenggara, Banten, Bali, Kalimantan Barat, DI Yogyakarta, Jawa Timur, Sumatra Utara, NTT, Jawa Barat, Jawa Tengah, Kalimantan Selatan, Bengkulu, Kalimantan Tengah, NTB, Maluku, DKI Jakarta, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Utara, Sumatra Barat, Jambi, Sumatra Selatan, sedangkan anggota *cluster* kedua antara lain Lampung, Gorontalo, Maluku Utara, Riau, Papua, Nanggroe Aceh Darussalam, dan Kalimantan Timur.

### 4.5.3 *Periodogram-based Distance*

Jarak *periodogram-based* diperoleh melalui perhitungan jarak *euclidean* dari nilai *periodogram* sebuah data deret waktu. DPK terlebih dahulu dibuat *periodogram*nya kemudian dihitung jarak *euclidean* dari *periodogram*-*periodogram* yang dihasilkan, kemudian dikelompokkan menggunakan analisis *cluster* hirarki. Jumlah kelompok yang dihasilkan menggunakan dapat dilihat melalui pemotongan dendrogram berdasarkan selisih jarak penggabungan terbesar pada Lampiran 9.

Selisih jarak penggabungan terbesar adalah selisih jarak penggabungan dari stage 29 dengan stage 28, yaitu sebesar 0,046201 sehingga jumlah kelompok yang terbentuk sebanyak 2 kelompok. Anggota tiap kelompok dapat diketahui melalui pemotongan dendrogram menjadi dua bagian. Pemotongan dendrogram dapat dilihat pada Gambar 4.5.



Gambar 4.5 Dendrogram Analisis *Cluster* Hirarki Menggunakan *Periodogram-based Distance*



Berdasarkan Gambar 4.6 jumlah *cluster* yang terbentuk sebanyak 2 *cluster* . Anggota kelompok 1 yaitu provinsi Nanggroe Aceh Darussalam, dan anggota kelompok 2 antara lain : Provinsi Bangka Belitung, Sulawesi Tenggara, Banten, Bali, Kalimantan Barat, DI Yogyakarta, Jawa Timur, Sumatra Utara, NTT, Jawa Barat, Jawa Tengah, Kalimantan Selatan, Bengkulu, Kalimantan Tengah, NTB, Maluku, DKI Jakarta, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Utara, Sumatra Barat, Jambi Sumatra Selatan, Lampung, Gorontalo, Maluku Utara, Riau, Papua, dan Kalimantan Timur.

#### 4.6 Ukuran Ketepatan Analisis *Cluster* Hirarki

*Cluster* yang terbentuk dari suatu metode pengelompokan mempunyai ukuran ketepatan hasil pengelompokan. Ukuran ketepatan dapat digunakan untuk mengetahui metode pengelompokan terbaik maupun jarak terbaik dengan membandingkan nilai ukuran ketepatan dari *cluster* yang dihasilkan. Ukuran ketepatan yang dipakai untuk mengetahui ketepatan pengelompokan dari jarak *autocorrelation-based*, *complexity invariant*, *periodogram-based*, dan *dynamic time warping* menggunakan analisis *cluster* secara hirarki yaitu dengan melihat kesesuaian pengelompokan dengan jarak pengelompokan, yang dapat dilihat melalui nilai korelasi *cophenetic* antara matriks jarak dengan matrik *cophenetic* pada tiap jarak yang dipakai.

Koefisien Korelasi *Cophenetic* digunakan untuk membandingkan dua Dendrogram, apakah dendrogram-dendrogram tersebut sudah sesuai dan dapat digunakan dengan mengetahui nilai korelasi antara jarak masing-masing objek dengan koefisien penggabungan masing-masing objek. Nilai korelasi *cophenetic* masing-masing jarak pengelompokan dapat dilihat pada Tabel 4.2. Pada Tabel 4.2 dapat diketahui bahwa jarak DTW merupakan jarak terbaik dilihat dari koefisien korelasi jarak DTW.

Tabel 4.2 Koefesien Korelasi *Cophenetic* Jarak *Autocorrelation-based*, *CID*, *Periodogram-based*, dan *DTW*

No	Jarak	Koefisien Korelasi <i>Cophenetic</i>
1	<i>Autocorrelation-based</i>	0,6731431
2	<i>Complexity Invariant</i>	0,8067845
3	<i>Periodogram-based</i>	0,7822575
4	<i>Dynamic Time Warping</i>	0,9775802

Berdasarkan Tabel 4.2 menunjukkan bahwa nilai korelasi *cophenetic* terbesar adalah nilai korelasi *cophenetic* jarak *Dynamic Time Warping* yaitu sebesar 0,9775802. Sehingga dapat diketahui jarak yang paling baik membentuk dendrogram adalah jarak *Dynamic Time Warping*, dibandingkan dengan jarak *Autocorrelation-based* yang mempunyai nilai korelasi *cophenetic* sebesar 0,6731431, *Complexity invariant distance* dengan nilai korelasi *cophenetic* sebesar 0,8067845, dan *Periodogram-based* yang mempunyai nilai korelasi *cophenetic* sebesar 0,7822575.

#### 4.7 Analisis Cluster Menggunakan Algoritma K-Means

Analisis *cluster* menggunakan algoritma *k-means* digunakan untuk mengoptimalkan hasil pengelompokan menggunakan analisis *cluster* hirarki. Berdasarkan hasil perhitungan korelasi *cophenetic* yang didapat diketahui bahwa pengelompokan terbaik adalah dengan menggunakan jarak *Dynamic Time Warping* (DTW). Hasil pengelompokan menggunakan jarak DTW yang diperoleh pada pemotongan dendrogram pada Gambar 4.6 menunjukkan Anggota kelompok 1 adalah provinsi Nanggroe Aceh Darussalam, sedangkan anggota kelompok 2 antara lain : Provinsi Bangka Belitung, Sulawesi Tenggara, Banten, Bali, Kalimantan Barat, DI Yogyakarta, Jawa Timur, Sumatra Utara, NTT, Jawa Barat, Jawa Tengah, Kalimantan Selatan, Bengkulu, Kalimantan Tengah, NTB, Maluku, DKI Jakarta, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Utara, Sumatra Barat, Jambi Sumatra Selatan, Lampung, Gorontalo, Maluku Utara, Riau, Papua, dan Kalimantan Timur.

Berdasarkan hasil pengelompokan melalui analisis *cluster* hirarki dapat diketahui dengan jelas kelemahan analisis *cluster* hirarki, yaitu dalam mengelompokan data yang memiliki pencilan. Provinsi Nanggroe Aceh Darussalam menjadi kelompok tersendiri dan provinsi lainnya menjadi satu kelompok. Oleh karena itu dibutuhkan optimalisasi pengelompokan menggunakan algoritma *k-means* dengan jumlah 2 kelompok dengan pusat kelompok 1 yaitu jumlah DPK pada Provinsi Nanggroe Aceh Darussalam, sedangkan pusat kelompok 2 adalah jumlah DPK rata-rata 29 provinsi pada kelompok 2. Anggota masing-masing *cluster* yang didapatkan melalui pengelompokan menggunakan algoritma *k-means* dapat dilihat pada Tabel 4.3.

Tabel 4.3 Anggota Tiap *Cluster* Hasil Analisis *Cluster* Menggunakan Algoritma *K-Means*

<i>Cluster</i>	Anggota <i>Cluster</i>
1	Nanggroe Aceh Darussalam, Riau, Bangka Belitung, Lampung, Kalimantan Timur, Gorontalo, Papua, Maluku Utara
2	Sulawesi Tenggara, Banten, Bali, Kalimantan Barat, DI Yogyakarta, Jawa Timur, Sumatra Utara, NTT, Jawa Barat, Jawa Tengah, Kalimantan Selatan, Bengkulu, Kalimantan Tengah, NTB, Maluku, DKI Jakarta, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Utara, Sumatra Barat, Jambi, dan Sumatra Selatan

Berdasarkan Tabel 4.3 dapat diketahui anggota masing-masing *cluster* yang terbentuk menggunakan analisis *cluster* dengan algoritma *k-means*. Menggunakan pusat *cluster* yang didapatkan dari analisis *cluster* hirarki, didapatkan Provinsi Nanggroe Aceh Darussalam dikelompokkan bersama dengan Provinsi Riau, Bangka Belitung, Lampung, Kalimantan Timur, Gorontalo, Papua, dan Maluku Utara. Sedangkan 22 provinsi lainnya yaitu Provinsi Sulawesi Tenggara, Banten, Bali, Kalimantan Barat, DI Yogyakarta, Jawa Timur, Sumatra Utara, NTT, Jawa Barat, Jawa Tengah, Kalimantan Selatan, Bengkulu, Kalimantan Tengah, NTB, Maluku, DKI Jakarta, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Utara, Sumatra Barat, Jambi, dan Sumatra Selatan dikelompokkan menjadi 1 *cluster*.

#### 4.8 Validasi Hasil Analisis *Cluster* dengan Algoritma *K-Means*

Analisis *cluster* menggunakan algoritma *k-means* menghasilkan dua *cluster* dengan anggota *cluster* pertama sebanyak 8 provinsi dan *cluster* kedua mempunyai anggota sebanyak 22 provinsi. Ketepatan pengelompokan menggunakan algoritma *k-means* dapat dilihat pada rata-rata koefisien *silhouette* dari masing-masing *cluster*. Kriteria penentuan ukuran ketepatan menggunakan koefisien *silhouette* dapat dilihat pada Tabel 2.3. Koefisien *silhouette* diperoleh dengan menghitung nilai rata-rata terbesar nilai *silhouette* pada jumlah *cluster* yang mungkin. Jumlah *cluster* yang mungkin dibentuk adalah sebanyak 28 ( $n-2$ ) jumlah kelompok yang mungkin. Nilai *silhouette* dari tiap jumlah

kelompok yang mungkin dapat dilihat pada Lampiran 11 . Koefisien *silhouette* pada Lampiran 11 menunjukkan nilai koefisien *silhouette* terbesar terdapat pada algoritma *k-means* dengan dua cluster yang terbentuk adalah yaitu sebesar 0,257.

Berdasarkan kriteria pada Tabel 2.3 dapat kita lihat bahwa nilai *silhouette* pada analisis *cluster* menggunakan jarak DTW dengan algoritma *k-means* termasuk dalam kriteria “*weak classified*” yang menunjukkan bahwa jumlah cluster yang terbentuk sudah dapat digunakan namun struktur yang dibentuk lemah, sehingga terdapat kemungkinan hasil pengelompokan tidak sesuai dengan kondisi sebenarnya.

#### **4.9 Akurasi Pengelompokan Jarak *Dynamic Time Warping***

Jarak *Dynamic Time Warping* bekerja secara langsung pada data deret waktu yang kompleks, berbeda dengan jarak atau algoritma yang lainnya, contohnya jarak *periodogram-based* yang bekerja pada periodogram yang didapatkan dari sebuah deret waktu, begitu juga jarak *autocorrelation-based* yang bekerja pada nilai fungsi autokorelasi yang dihasilkan oleh sebuah deret waktu. Cara kerja algoritma DTW secara langsung pada data deret waktu memungkinkan *distorsi* (pergeseran) yang sangat kecil pada Dendrogram yang dihasilkan dengan nilai kedekatan sesungguhnya antar dua deret waktu. Hal tersebut dapat dilihat melalui nilai korelasi *cophenetic* jarak DTW yaitu sebesar 0,9775802. Besarnya nilai korelasi *cophenetic* menunjukkan kesesuaian antara matriks jarak DTW dengan matriks penggabungan *cluster* hirarki, sehingga semakin besar nilai korelasi *cophenetic* maka semakin kecil pula *distorsi* antara kedekatan sesungguhnya dari dua deret waktu dengan Dendrogram yang terbentuk melalui analisis *cluster* hirarki.

Jarak DTW menghasilkan struktur pengelompokan yang kurang baik pada optimalisasi analisis *cluster* secara non-hirarki. Berdasarkan nilai koefisien *silhouette* yang didapat sebesar 0,257 menunjukkan bahwa jarak DTW mampu menghasilkan kelompok yang dapat diterima namun struktur pengelompokannya lemah sehingga dikhawatirkan terjadi bias antara hasil analisis *cluster* dengan kondisi sesungguhnya.

UNIVERSITAS BRAWIJAYA



## BAB V

### KESIMPULAN

#### 5.1 Kesimpulan

1. Penggunaan jarak *Autocorrelation-based*, *Complexity Invariant*, *Periodogram-based*, dan *Dynamic Time Warping* memiliki hasil pengelompokan yang berbeda.
2. Jarak *Dynamic Time Warping* memiliki koefisien korelasi *cophenetic* terbesar yaitu sebesar 0,9775802, sehingga Jarak *Dynamic Time Warping* memiliki akurasi pengelompokan terbaik dibandingkan jarak *Autocorrelation-based*, *Complexity Invariant*, *Periodogram-based*. Jarak *Dynamic Time Warping* (DTW) dengan nilai koefisien silhouette sebesar 0,257 menghasilkan hasil pengelompokan yang cukup baik pada analisis *cluster* dengan algoritma *K-means*.
3. Pengelompokan data deret waktu Dana Pihak Ketiga Menghasilkan dua kelompok. Anggota kelompok pertama adalah Nanggroe Aceh Darusslam, Bangka Belitung, Gorontalo, Maluku Utara, Riau, Papua, Kalimantan Timur, dan Lampung . Anggota kelompok dua antara lain Provinsi Sulawesi Tenggara, Banten, Bali, Kalimantan Barat, DI Yogyakarta, Jawa Timur, Sumatra Utara, NTT, Jawa Barat, Jawa Tengah, Kalimantan Selatan, Bengkulu, Kalimantan Tengah, NTB, Maluku, DKI Jakarta, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Utara, Sumatra Barat, Jambi, dan Sumatra Selatan.

#### 5.2 Saran

Analisis *cluster* data deret waktu merupakan analisis yang sangat kompleks dan merupakan salah satu aplikasi dari *data mining* sehingga diperlukan kecermatan dalam melakukan pengelompokan. Dalam membandingkan akurasi dari sebuah pengelompokan dapat menggunakan berbagai macam ukuran ketepatan sehingga ketepatan pengelompokan dapat diketahui dengan baik.

UNIVERSITAS BRAWIJAYA



## DAFTAR PUSTAKA

- Ahman, E. dan Indriani, E. 2007. *Membina Kompetensi Ekonomi*. Bandung : Grafindo Media Pratama.
- Batista, G. E., Wang, X., dan Keogh, E. J. 2011. “*A Complexity-Invariant Distance Measure for Time Series*”. In *SDM* (Vol. 31, p. 32).
- Box, G.E.P, Jenkins, G.M, dan Reinsel, G.C.2008.*Time Series Analysis Forecasting and Control*.Fourth Edition.New Jersey: John Wiley and Sons Inc.
- Caiado, J., Catro, N., and Pena, D. 2006. A Peridogram-based Metric for Time Series Clasification. *Computational Statistics & Data Analysis*.50.2668 – 2684.
- Charles. 2006. *Analisis Perbandingan Pemodelan Data Deret Waktu Terbaik Antara Metode Brown’s Double Exponential Smoothing, Hold’s Two Parameter Trend Model dan ARIMA pada Total Penjualan Produk Optik Berbasiskan Komputer*. Universitas Bina Nusantara. Jakarta.
- Casado de Lucas, D. 2010. *Classification techniques for time series and functional data*.Disertasi Doktor pada Universitas Carlos III de Madrid.
- Cryer, J.D and Chan, K. 2008. *Time Series Analysis : with Aplication in R*.Second Edition.New york : Springer+business media.
- Dendawijaya, L. 2005. *Manajemen Perbankan*. Edisi Kedua, Cetakan Kedua. Bogor : Ghalia Indonesia.

Everitt, B.S., Sabine, L., Leese, M., dan Stahl, D.2011.*Cluster Analysis*.Fifth Edition.United Kingdom: JohnWilly and Son,Ltd.

Fernandes, A.R. 2008. *Modul Kuliah Eksplorasi Data Peubah Ganda*. Prodi Statistika, Universitas Brawijaya, Malang (tidak dipublikasikan).

Galeano,P dan Pena, D.2000."Multivariate Analysis in Vector Time Series". Resenhas,4(4):383-403. ISSN 0104-3854.

Hair, J.F., Black, W.C., Babin, B.J. 2010. *Multivariate Data Analysis*. Seventh Ed.New Jersey: Prentice Hall International, Inc.

Han, J. and Kamber, M. 2012. *Data Mining and Techniques.Third Ed*. New York : Elsevier,inc.

Harris, R. and Sollis, R. 2003. *Applied Time Series Modelling and Forecasting*. New York: Wiley

Kaufman, L. dan Rousseuw, P.J. 1990. *Finding Group in Data : An Introduction to Cluster Analysis*. New York: Wiley

Liao,T.W. 2005. Clustering of Time Series Data—Survey.*Pattern Recognition*. Vol.38. Page : 1857–1874.

Maharaj,E.A. 2000. Clusters of time series. *J. Classification*. Vol.17. Page : 297–314

Makridakis, S., Steven, C.W., and Victor E.1999. *Forecasting, 2nd Edition*. New York : John Wiley and Sons, Inc.

Manly. 1988. *Multivariate Statistical Methods*. New York : A Primer Champma an Hall.

- Manso, P.M. 2013. *A Package for Stationary Time Series Clustering*. Tesis Master pada Universitas da Coruna, Spanyol.
- Marrison, D.F. 1978. *Multivariate Statistical Methods*. New York : Mc Graw Hall.
- Moller-Levet, C.S., Klawonn, F., Cho, K.H., Yin, H., Wolkenhauer, O. 2005. Clustering of unevenly sampled gene expression time-series data. *Fuzzy Sets and Systems*. Vol.152. Page : 49-66.
- Montgomery, D.C, Jennings, C.L., dan Kulahci, M. 2008. *Introduction to Time Series Analysis and Forecasting*. New York : John Willey & Sons, Inc.
- Niennattrakul, V. dan Ratanamahatana, C.A. 2007. On Clustering Multimedia Time Series Data Using K-Means and Dynamic Time Warping. *International Conference Multimedia and Ubiquitous Engineering*. Vol.07. Page : 733-738.
- Piccolo, D. 1990. A distance measure for classifying ARMA models. *J. Time Series Analysis*. Vol.11. No.2. Page : 153-163.
- Rani, Sangeeta., Geeta, S. 2012. "Recent Techniques of Clustering of Time Series Data: a Survey". *International Journal of Computer Application*. Vol.52. No.15.
- Sakoe, H dan Chiba, S. 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*. Vol.26. Page : 43-49.
- Senin, P. 2008. *Dynamic Time Warping Algoritma Review*. Information and Computer Science Department University of Hawaii at Manoa.

Suhardi, G. 2003. *Usaha perbankan dalam aspek hukum*. Jakarta: kanisius.

Suprianto, J. 2008. *Statistik: Teori dan Aplikasi, Jilid 2*. Jakarta : Erlangga

Usman, R. 2003. *Aspek - Aspek Hukum Perbankan Indonesia*. Jakarta: PT.Garamedia Pustaka Utama.

Wei,W.W.S. 2006. *Time Series Analysis Univariate and Multivariate Methods*. Second Edition. United State of America : Pearson Education, Inc.



Lampiran 1. Data dana Pihak Ketiga Januari 2005 - Desember 2013 30 provinsi di Indonesia

1. Dana Pihak Ketiga Tahun 2005

Provinsi	Jan	Feb	Maret	April	Mai	Juni	Juli	Agst	Sept	Oktr	Nov	Drs
Jawa Barat	75304	74731	74838	73894	75778	79083	80716	81256	82891	81759	84647	87899
Banten	18877	19038	19098	19085	19239	19587	19263	19832	20390	21053	21822	21706
DKI	498624	495897	505043	517538	523241	529677	537674	554204	571827	566642	572945	586432
Jakarta	498624	495897	505043	517538	523241	529677	537674	554204	571827	566642	572945	586432
DI	498624	495897	505043	517538	523241	529677	537674	554204	571827	566642	572945	586432
Yogyakarta	10249	10199	10317	10447	10565	10805	10672	10907	11153	11072	11131	11453
Jawa	45367	45568	46129	47404	47445	48584	48265	49594	50801	50178	51603	52950
Tengah	45367	45568	46129	47404	47445	48584	48265	49594	50801	50178	51603	52950
Jawa	93969	94484	94676	97869	99392	101425	100451	103714	108155	108436	110324	111909
Timur	93969	94484	94676	97869	99392	101425	100451	103714	108155	108436	110324	111909
Bergabung	1697	1690	1679	1649	1652	1784	1766	1816	1821	1808	1825	2066
Jambi	5063	5105	5110	5127	5140	5364	5325	5447	5425	5420	5524	5707
Aceh	7702	8468	8229	8124	8118	9321	9482	10208	10975	9851	10666	13511
Sumatra Utara	43855	43821	44684	45417	44891	45979	45593	46361	48876	49014	49979	51589
Sumatra Barat	7802	7760	7731	7779	7795	8058	7963	8194	8288	8260	8620	9109
Riau	22697	22792	22801	22987	22649	24739	24257	25761	25383	24631	26355	28558
Sumatra Selatan	14452	14457	13951	14369	14476	15090	14879	15442	15503	15789	16328	17848
Bangka Belitung	2991	2960	2838	2941	2998	3184	3215	3189	3321	3315	3401	3497
Lampung	6272	6209	6229	6327	6289	6833	6686	7025	7247	6876	6997	7419

## Lampiran 1. Lanjutan

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Okt	Nov	Des
Kalimantar Selatan	6688	6717	6703	6740	6811	7170	7104	7381	7552	7577	7644	8122
Kalimantar Barat	8566	8568	8669	8856	8862	9168	9141	9389	9674	9629	9759	10291
Kalimantar Timur	17228	17169	17061	17087	16992	19014	18583	20415	19935	21294	21239	23281
Kalimantar Tengah	3248	3360	3470	3469	3483	3682	3586	3695	3739	3692	3753	4109
Sulawesi Tengah	2721	2702	2809	2854	2885	3177	3000	3013	3073	2964	3008	3313
Sulawesi Selatan	14142	14202	14390	14445	14561	15325	14972	15265	16097	16043	16171	17239
Sulawesi Utara	4254	4196	4382	4296	4284	4474	4327	4438	4716	4446	4475	4862
Gorontalo	778	759	750	732	745	823	824	838	866	820	877	1011
Sulawesi Tenggara	1939	1830	1898	1883	1873	2032	1969	1980	2142	2003	2079	2202
NTB	3180	3173	3131	3183	3172	3336	3276	3304	3617	3546	3593	3770
Bali	15252	15159	15325	15628	15701	16072	16308	16968	17263	17196	17142	17186
NTT	4217	4130	4167	4192	4147	4304	4305	4462	4456	4567	4520	4910
Maluku	2715	2654	2644	2656	2610	2599	2624	2559	2687	2770	2816	3248
Papua	6800	6484	6207	6206	6083	6346	6561	6935	6999	7265	8262	8857
Maluku Utara	768	739	721	711	725	750	720	688	754	801	813	1025

Lampiran 1. Lanjutan

2. Dana Pihak Ketiga Tahun 2006

Provinsi	Jan	Feb	Maret	April	Mai	Juni	Juli	Agst	Sept	Oktr	Nov	Des
Jawa Barat	86032	88786	88787	90263	92911	93941	93948	92860	97254	99293	103292	104431
Banten	22709	22136	22591	22947	23507	24040	24051	23973	24609	25997	26274	27317
DKI	566609	574487	578165	583676	574160	577762	571458	579605	591381	607615	612465	620919
D.I												
Yogyakarta	11552	11557	11651	11738	12117	12498	12648	13036	12996	13045	13163	13908
Jawa Tengah	53926	54482	54491	55254	56565	56998	56935	58668	58275	58824	59943	62701
Jawa Timur	115184	114978	115381	115546	118851	119666	118990	121720	121657	122720	124716	127851
Bengkulu	2142	2197	2267	2327	2404	2435	2437	2596	2624	2689	2889	2760
Jambi	5775	5907	5886	6062	6493	6432	6521	6826	6825	7009	7151	7335
Aceh	13222	13293	13357	13800	15633	15476	15825	17383	17070	17422	18033	21955
Sumatra Utara												
Lhara	51378	51295	52144	52327	52974	53917	54241	55898	57304	57688	58376	60125
Sumatra Barat	8858	8870	9152	9352	9725	9792	9915	10291	10570	10597	11005	11490
Riau	28571	28381	28292	28225	33565	33535	32237	35802	35575	36844	37838	40074
Sumatra Selatan	17660	17578	17654	17631	18932	18814	18791	20120	19783	20454	20528	21068
Bangka Belitung	3710	3773	3905	3968	4209	4327	4255	4429	4613	4615	4541	4728
Lampung	7721	7706	7922	8057	8608	8679	8655	8927	9050	9245	9592	9495

Lampiran 1. Lanjutan

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Okt	Nov	Des
Kalimantan Selatan	8466	8501	8619	8894	9347	9533	9481	9672	9577	9844	10306	10813
Kalimantan Barat	10578	10906	10872	11121	11562	11748	11763	12041	12128	12329	12542	12793
Kalimantan Timur	23216	22876	22522	22109	26167	25876	25152	27455	26951	28721	28870	30596
Kalimantan Tengah	4192	4274	4472	4776	4878	5161	5166	5294	5275	5271	5385	5510
Sulawesi Tengah	3275	3370	3485	3588	3786	3902	3933	4004	4056	4081	4206	4448
Sulawesi Selatan	17411	17336	17464	17661	18951	18876	18825	19114	19399	19959	20281	21952
Sulawesi Utara	5058	5014	5093	5079	5273	5330	5308	5420	5478	5444	5448	6057
Gorontalo	1024	1057	1071	1138	1229	1264	1233	1271	1279	1291	1293	1392
Sulawesi Tenggara	2455	2431	2615	2717	2869	2968	2880	2997	3066	3073	3341	3218
NTE	4019	4003	4056	4164	4385	4358	4291	4293	4387	4551	4205	4441
Bali	17350	17292	17243	17311	17716	17866	17888	18104	18172	18390	18675	19030
NTI	4592	5126	5186	5397	5666	5688	5769	5987	6003	6179	6160	6228
Maluku	2885	2899	2981	3034	3222	3265	3316	3420	3450	3528	3396	3993
Papua	9217	9327	9887	10343	10937	11068	11255	12782	12794	13002	12996	15096
Maluku Utara	1581	1502	1607	1806	1754	1750	1694	1753	1808	1754	1827	2184

## Lampiran 1. Lanjutan

## 3. Dana Pihak Ketiga Tahun 2007

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Oktr	Nov	Des
Jawa Barat	103961	104592	103102	105396	105566	107248	110931	107414	107880	113962	113870	119504
Banten	25925	25532	26296	26208	26780	26538	26932	27334	27173	27515	27714	28310
DKI												
Jakarta	612728	613851	619880	621421	627898	600506	670219	676546	682107	689257	697984	751446
DI												
Yogyakarta	13979	13960	13884	14047	13996	14121	14221	14769	14562	14797	14982	15323
Jawa Tengah	62227	62941	63519	64482	64485	65277	66207	67245	67759	69207	69100	70159
Jawa Timur	127747	129521	126612	129975	130144	133450	136078	137591	137281	137618	139384	143778
Bengkulu	2775	2826	2901	2944	3050	3196	3298	3458	3536	3580	3622	3482
Sumba	7542	7642	7703	7938	7952	8180	8480	8858	8727	9046	9256	9321
Sumatra Utara	20946	20165	20041	19763	19251	19293	19456	19819	19275	18390	18026	18341
Sumatra Utara	59300	59590	60777	61072	60698	62329	63542	65004	65947	68334	69531	71021
Sumatra Barat	11036	11326	11485	11663	12050	12198	12424	12685	12819	13388	13946	13674
Riau	26829	26471	26176	26274	25980	27278	27758	29285	28858	28743	30079	30155
Sumatra Selatan	21920	21195	21152	20983	20641	21427	21696	22531	22986	23119	23800	24867
Bangka Belitung	4679	4707	4776	5015	5064	5451	5486	5680	5015	5331	6611	6384
Lampung	15958	16099	16106	16502	16168	15820	15778	15468	14526	14494	14383	14789

Lampiran 1. Lanjutan

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Oktr	Nov	Des
Kalimantan Selatan	10603	10770	11041	11086	11356	11651	11656	11963	12282	12358	12674	12798
Kalimantan Barat	12832	13136	13354	13692	13576	13791	14139	14354	14498	14678	14818	14965
Kalimantan Timur	30457	30611	30506	30125	29228	30138	30607	32300	32490	32117	33695	34316
Kalimantan Tengah	5228	5324	5671	5798	5993	6051	6221	6287	6496	6641	6708	6530
Sulawesi Tengah	4584	4480	4502	4575	4640	4787	4798	4949	4984	4927	5073	5171
Sulawesi Selatan	20005	20080	20122	20459	20883	22031	23622	22577	22893	23442	23607	24794
Sulawesi Utara	5948	5932	6032	6091	6261	6484	6556	6596	6554	6571	6605	7133
Gorontalo	2004	2027	2035	2078	2125	1566	1566	1574	1558	1584	1577	1641
Sulawesi Tenggara	3450	3450	3486	3551	3510	3671	3731	3818	3903	3857	3977	3977
NTB	4827	4928	5048	4952	4966	5040	5125	5107	5210	5327	5424	5402
Bali	19939	19594	19704	19584	20356	20671	21142	21669	21908	22516	23084	23522
NTT	6459	6559	6654	6773	6891	6933	7103	7217	7141	7185	7237	7296
Mahuku	3883	3945	4033	3817	3832	3975	4077	4153	4173	4210	4250	4413
Papua	11199	11114	11114	10836	10745	10821	12143	12271	12356	12090	12250	13276
Mahuku Utara	2148	2082	2148	2094	2155	2210	2281	2293	2290	2351	2392	2576

Lampiran 1. Lanjutan

4. Dana Pihak Ketiga Tahun 2008

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Oktr	Nov	Des
Jawa Barat	115861	116395	115929	117400	119447	121288	121552	120619	123094	130429	138817	135580
Banten	27956	28171	28998	30592	30431	30332	28396	28315	28935	29945	30283	33011
DKI	729352	726595	717001	723572	733061	765022	745106	736585	785919	819032	840976	838803
Jakarta												
DI												
Jogyakarta	15545	15545	15471	15693	15590	15774	15860	15888	16103	16722	16838	16834
Jawa Tengah	68578	69555	69966	70695	71575	73784	74026	74309	76202	79516	80062	80782
Jawa Timur	141837	143560	142926	145219	147564	150226	148840	150127	155232	166221	168001	171426
Bergabung	3511	3554	3722	3706	3772	4007	4009	3919	4363	4478	4418	4143
Jambi	9245	9456	9587	9790	10101	10361	10458	10141	10342	10666	10433	10069
Acet	17176	16732	16595	16591	16850	17062	17529	17551	19177	18641	19526	20454
Sumatra Utara												
Jawa	70391	72123	71746	72965	73940	75416	76045	76526	77633	82290	84295	83941
Sumatra Barat	13486	14193	14018	13834	13810	13972	14061	13991	14017	15309	15372	15349
Riau	29491	28991	28372	28613	29908	30048	30445	29260	31953	31626	30933	31760
Sumatra Selatan	24612	24670	23926	24332	24458	25805	26311	25482	26873	28934	29518	30058
Bangka Belitung	6284	6554	6487	6822	6769	7374	7066	7064	7207	7429	7284	6328
Lampung	14036	13367	13216	13122	13477	14387	14350	14536	15559	16505	16839	21559

Lampiran 1. Lanjutan

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Oktr	Nov	Des
Kalimantan Selatan	12572	12994	13323	13622	14068	14414	14531	14752	15412	15573	15699	16026
Kalimantan Barat	14738	15558	15394	15627	15529	15913	16076	16014	16360	16952	17070	17570
Kalimantan Timur	33529	33004	32974	32392	34506	35113	36530	35997	39350	39353	39334	41518
Kalimantan Tengah	6559	6738	6962	7102	7264	7419	7577	7287	7611	7712	7587	6909
Sulawesi Tengah	4839	5053	5137	5153	5238	5582	5357	5237	5320	5755	5715	5721
Sulawesi Selatan	24063	24225	24388	24476	25133	26231	26131	26822	26737	27524	28017	29013
Sulawesi Utara	6948	7119	7263	7081	7260	7841	7575	7717	8004	8376	8510	8941
Gorontalo	1503	1564	1595	1595	1570	1661	1610	1567	1700	1799	1787	1833
Sulawesi Tenggara	3934	4063	4092	4153	4156	4292	4323	4240	4492	4808	4927	4563
NTB	5246	5339	5361	5434	5468	5513	5599	5527	6023	6297	6497	6363
Bali	23550	24006	24267	24754	25298	26675	25931	26192	26576	27816	28540	28524
NTT	6987	7087	7186	7210	7307	7438	7295	7312	7887	7935	7963	8005
Maluku	4319	4439	4338	4331	4367	4396	4556	4341	4569	4655	4703	4671
Papua	12512	12464	12112	11963	12570	12810	12579	13335	15359	15088	15065	14227
Maluku Utara	2526	2662	2667	2636	2635	2692	2598	2462	2656	2826	2830	2800

## Lampiran 1 Lanjutan

## 5. Dana Pihak Ketiga Tahun 2009

Provinsi	Jan	Feb	Maret	April	Mei	Jun	Juli	Agst	Sept	Okt	Nov	Des
Jawa Barat	135624	137019	141274	138885	139721	143452	142748	145732	148445	149443	149457	154212
Banten	32945	33050	33303	34108	31132	34457	33531	33383	34990	34790	35912	39103
DKI	865982	881885	880839	872078	879948	899351	893947	923663	921935	920423	948887	995417
DI	17103	17201	17502	17654	17625	18039	18285	18846	19152	19262	19362	19679
Yogyakarta	17103	17201	17502	17654	17625	18039	18285	18846	19152	19262	19362	19679
Jawa	81287	82294	85453	84765	84764	86479	85817	85874	87621	88278	88803	91223
Tengah	81287	82294	85453	84765	84764	86479	85817	85874	87621	88278	88803	91223
Jawa	172528	176088	177520	178460	180061	181895	182692	183183	186079	187554	189786	193720
Timur	172528	176088	177520	178460	180061	181895	182692	183183	186079	187554	189786	193720
Bengkulu	4144	4106	4192	4249	4257	4203	4093	4123	4053	4071	4081	3967
Lambi	9937	10044	10281	10484	10388	10564	10292	10384	10231	10274	10452	10843
Acch	19034	18104	18018	18746	18335	18379	17441	18515	18450	18075	18834	19215
Sunatra	85123	86979	88454	88095	87643	89156	88843	90211	89906	91392	91781	94458
Utara	85123	86979	88454	88095	87643	89156	88843	90211	89906	91392	91781	94458
Sunatra	15432	15693	16228	16479	16273	16297	16195	16171	16065	16594	16853	17728
Barat	15432	15693	16228	16479	16273	16297	16195	16171	16065	16594	16853	17728
Riau	31325	31316	31710	32154	31521	33072	32158	32131	31642	31485	30548	30892
Sunatra	29249	29644	30374	29763	29604	30102	28950	29424	29874	30027	30881	34119
Selatan	29249	29644	30374	29763	29604	30102	28950	29424	29874	30027	30881	34119
Bangka	6842	7011	7477	7783	7721	7625	7377	7375	7592	7606	7558	7469
Belitung	6842	7011	7477	7783	7721	7625	7377	7375	7592	7606	7558	7469
Lampung	21508	19539	17684	17684	17718	13785	13378	13746	13977	14157	14300	14836

Lampiran 1. Lanjutan

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Okt	Nov	Des
Kalimantan Selatan	16127	16292	17154	17060	16774	17194	16838	16850	17216	17250	17312	18164
Kalimantan Barat	17560	17329	18125	17807	18240	18412	18405	18907	18789	19058	18660	18995
Kalimantan Timur	40180	40221	41593	43236	40389	42379	41215	41188	42954	42576	41879	43719
Kalimantan Tengah	7221	7349	7621	7823	7845	8115	8114	8072	7951	8150	8020	7708
Sulawesi Tengah	5565	5526	5705	5842	5927	6091	5854	5836	5960	6025	5063	6301
Sulawesi Selatan	28375	28659	29389	29294	29580	30152	29672	29724	30771	30853	31535	34271
Sulawesi Utara	8683	8647	8992	9237	9279	9529	9445	9504	9806	9885	9836	10082
Gorontalo	1857	1866	1923	1974	1924	1947	1890	1877	1952	1954	1929	1909
Sulawesi Tenggara	4830	4826	4861	5033	4863	5093	4953	5037	5010	5142	5079	5207
NTB	6517	6469	6613	6759	6745	6809	6579	6801	7005	7078	7153	7101
Bali	29032	29352	29565	29403	29408	30254	30321	31145	31364	31622	31863	32247
NTT	8055	8019	8269	8626	8742	8824	8775	8911	8940	9092	9194	9119
Maluku	4731	4735	4799	4872	4891	4878	4925	4971	5044	5148	5089	5101
Papua	13834	13891	14331	14876	14810	16453	15807	15761	16179	17006	16667	16344
Maluku Utara	2809	2688	2828	2751	2732	2896	2685	2554	2640	2727	2709	2673

Lampiran 1. Lanjutan

6. Dana Pihak Ketiga Tahun 2010

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Okt	Nov	Des
Jawa Barat	151892	151957	153738	154650	158952	164822	167078	166682	170138	175575	177719	186358
Banten	39185	38056	41191	43962	44108	47445	46530	45994	46631	49655	54273	56531
DKI												
Jakarta	982294	967783	994088	993045	1012718	1057986	1042332	1054018	1075943	1094360	1115796	11197604
DI												
Yogyakarta	19772	19820	20122	20263	20558	21119	21521	21224	21464	21734	22309	22919
Jawa Tengah	90940	90571	92869	93573	94843	98029	97725	97428	100554	100552	101166	105605
Jawa Timur	191172	190364	192713	192040	195383	199013	199445	199006	202293	203256	207935	214993
Bengkulu	4082	4079	4164	4192	4284	4453	4622	4614	4643	4688	4718	4691
Jambi	10750	10787	11471	11340	11463	12161	11753	11595	12440	12439	12550	13254
Aceh	17444	16775	16873	16377	17101	17399	16762	17134	18952	18129	17716	18851
Sumatra Utara	94218	93945	94963	93546	94267	98546	98805	100365	102485	103951	105506	108612
Sumatra Barat	17820	17644	17595	17492	17462	17858	18344	17985	18569	18711	18952	20829
Riau	29874	30044	33886	33549	33843	34337	33680	33105	35424	34960	34598	37142
Sumatra Selatan	33360	33213	35293	35148	35030	35788	36791	36875	37835	38218	38779	40559
Bangka Belitung	7617	7638	8004	8200	8188	8264	8104	8142	8293	8294	8538	8645
Lampung	14591	14511	14732	14579	14862	15722	15709	15893	16122	15861	16143	16748

Lampiran 1. Lanjutan

Provinsi	Jan	Feb	Maret	April	Mei	Jun	Juli	Agst	Sept	Okt	Nov	Des
Kalimantan Selatan	18113	18109	18664	18682	18773	19605	19602	19549	20327	20248	20695	21572
Kalimantan Barat	19031	18860	19524	19868	20184	20829	20921	21002	21562	21728	22287	23071
Kalimantan Timur	42205	42571	46590	45176	44621	47995	47001	46103	49366	48650	47573	49911
Kalimantan Tengah	8013	8090	8379	8459	8470	8956	8903	8768	9139	9005	8935	8899
Sulawesi Tengah	6315	6313	6414	6273	6351	5891	6843	6884	7018	7018	7154	7450
Sulawesi Selatan	32859	32537	32385	31918	32387	33838	34408	34486	35081	35219	36028	38038
Sulawesi Utara	10020	10095	10324	10354	10352	10694	10854	11055	11218	11038	11347	11554
Gorontalo	1927	1958	1928	1944	1939	2101	2133	2077	2183	2171	2154	2191
Sulawesi Tenggara	5336	5265	5250	5258	5242	5652	5676	5627	5769	5935	5958	5972
NTB	7165	7143	7261	7282	7433	7797	7685	7800	7904	8110	8087	8473
Bali	32825	32454	32542	32555	32717	33649	34336	35029	35734	36500	36962	37848
NTT	9255	9306	9231	9211	9391	9767	9844	9772	9829	9737	10373	10208
Mahuku	5236	5134	5143	5052	5037	5318	5412	5360	5440	5403	5406	5710
Papua	16317	16210	16454	16382	17102	18678	18470	18417	18761	19885	19548	19334
Mahuku Utara	2795	2638	2695	2632	2604	2743	2760	2649	2765	2773	2720	2750

Lampiran 1 Lanjutan

7. Dana Pihak Ketiga Tahun 2011

Provinsi	Jan	Feb	Marci	April	Mei	Juni	Juli	Agst	Sept	Oktr	Nov	Des
Jawa Barat	183504	183688	185128	185745	188931	190846	196230	197080	204444	205757	210565	215405
Banten	55351	55882	55183	58921	60305	62217	63126	63585	65972	65734	68953	70218
DKI	1166975	1141145	1180064	1164689	1197538	1202544	1225588	1230557	1255384	1293125	1320964	1417286
Jakarta												
DI												
Jogyakarta	23186	23005	23276	23577	23981	24323	24840	24776	25839	26134	25270	26837
Jawa Tengah	104882	105040	109546	110668	112895	115121	116897	111322	118301	118404	122107	124743
Jawa Timur	213054	213282	216775	217690	221755	225753	228094	225597	234007	235536	241240	250524
Bengkulu	4944	5035	5188	5316	5482	5686	5752	5940	5986	6024	6232	6190
Lambi	13606	14338	14478	14639	14857	15437	15440	15335	16079	15973	15149	16415
Aceh	18030	18106	17922	17887	19275	19295	18554	17956	20063	19438	19251	20333
Sumatra Utara	109453	110086	112139	112630	114365	115920	117380	115577	120200	120910	124317	127076
Sumatra Barat	20503	20204	20206	20461	20835	20960	21164	20704	22554	22319	23076	23126
Riau	36322	39149	40188	39073	40114	42534	41856	41415	44115	43074	43189	45057
Sumatra Selatan	40373	41661	42744	43609	44695	46538	45517	46887	48668	49377	50378	51235
Bangka Belitung	8739	9004	9523	9448	9727	10277	10458	10235	10616	10230	10129	10243
Lampung	16706	16801	17454	17563	17988	18566	18568	18354	19272	19150	19561	19745

Lampiran 1. Lanjutan

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Okt	Nov	Des
Kalimantan Selatan	21550	21579	22383	22447	23158	24285	24442	24374	26282	26262	27006	28305
Kalimantan Barat	23435	23720	24264	24641	25089	25324	25338	25481	26213	26327	27597	27857
Kalimantan Timur	49254	52444	53331	51586	53126	56831	56446	57158	61695	61829	62426	66880
Kalimantan Tengah	9233	9453	9862	10631	11090	11505	11394	11456	12064	12099	12003	11924
Sulawesi Tengah	7466	7430	7464	7425	7686	8003	7818	7767	8054	8241	8593	8865
Sulawesi Selatan	37519	37432	38224	38170	38688	39908	40205	40088	41879	42350	43727	46252
Sulawesi Utara	11702	11659	11929	12059	12341	12738	12806	12802	13447	13563	13908	14340
Gorontalo	2278	2305	2343	2353	2416	2425	2411	2389	2550	2555	2738	2712
Sulawesi Tenggara	6365	6407	6583	6787	7006	7165	7170	7479	7962	8155	8256	8428
NTB	8584	8445	8564	8872	9176	9379	9574	9298	10018	10246	10594	10872
Bali	38356	38213	38536	38854	39542	40430	40966	41124	42801	43946	44739	45604
NTT	10651	10684	10791	11021	11292	11423	11537	11522	11877	11974	12438	12755
Mahuku	5650	5777	5909	5008	6096	6215	6131	6126	6409	6477	6735	6989
Papua	15902	20004	21515	21584	21174	21704	21391	20564	23995	23945	23857	23546
Mahuku Utara	2978	3032	3123	3187	3184	3274	3258	3194	3431	3552	3714	3769

Lampiran 1. Lanjutan

8. Dana Pihak Ketiga Tahun 2012

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Ok	Nov	Des
Jawa Barat	217474	217502	225244	229017	236025	237542	235355	242524	244432	247308	254557	252433
Banten	70304	72345	74852	77570	78552	79028	79845	81717	81436	82856	82298	88486
Jakarta	1410213	1385540	1410500	1407328	1458755	1477968	1480161	1488556	1510724	1524850	1566877	629585
Jogyakarta	25985	27485	28029	28488	28576	29227	29723	30425	31026	30922	31719	32501
Jawa Tengah	124768	125994	129494	131524	132133	134957	134838	137370	141489	140417	142949	145257
Jawa Timur	248576	250120	255493	256319	250011	255056	257058	258903	277597	280955	283615	292224
Bengkulu	6380	6361	6616	6770	6554	7108	7287	7353	7494	7614	7522	7358
Jambi	15313	15562	17255	17435	17504	17612	17506	17208	17918	17725	17707	17945
Aceh	19505	19526	19706	20594	20693	20558	20015	19348	21414	21551	21140	21959
Sumatra Utara	125631	127559	128448	128817	129117	129083	131537	133098	135506	135990	136397	139259
Sumatra Barat	22344	22992	23593	23628	23818	24251	24538	24715	25852	25503	25829	25623
Riau	44355	45567	48480	48445	48653	50314	50974	50375	53457	52553	51749	52243
Sumatra Selatan	50925	51718	53280	53322	53287	52071	52173	51451	52833	52763	52755	54244
Bangka Belitung	10335	10691	10849	10812	10910	10920	10715	10758	11877	12242	12303	12172
Lampung	19571	19607	20474	21156	21451	22056	22008	21565	22502	22889	22958	22612
Kalimantan Selatan	28216	28935	29547	30237	30978	31881	31257	31209	32836	32858	33521	34118

Lampiran 1. Lanjutan

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Okt	Nov	Des
Kalimantan Barat	27989	28371	28856	29038	29572	30352	30270	30419	31050	30696	31882	32000
Kalimantan Timur	65287	69242	73171	79209	72640	77217	76515	75582	79990	79022	78155	79992
Kalimantan Tengah	12396	12897	13458	13848	14146	14552	14208	14272	14785	15526	15480	14235
Sulawesi Tengah	8878	8911	9358	9603	9765	10064	9904	9542	10177	10070	10106	10022
Sulawesi Selatan	45591	45556	46249	47307	47604	48525	48552	49435	51079	51502	52445	54453
Sulawesi Utara	14358	14409	14785	14987	15069	15384	15346	15328	15795	15852	15991	16407
Gorontalo	2787	2751	2874	2944	2916	2995	2906	2828	3027	3005	3008	2990
Sulawesi Tenggara	8580	8648	9258	9617	9701	9930	9728	9532	10020	9933	10081	9870
NTB	10306	10753	11005	11259	11383	11900	11735	11752	12355	12495	12525	12734
Bali	45807	45295	45898	47413	48101	49577	50254	51492	52938	53883	54555	54948
NTT	12977	12916	13450	13814	13893	14296	14477	14457	14752	14698	14834	14834
Maluku	7216	7229	7199	7383	7455	7684	7630	7752	7894	8022	8233	8340
Papua	23970	24159	24352	25529	25798	26298	25771	27739	27932	27506	27054	27148
Maluku Utara	3892	4031	4313	4238	4170	4258	4273	4220	4452	4456	4493	4425

## Lampiran 1. Lanjutan

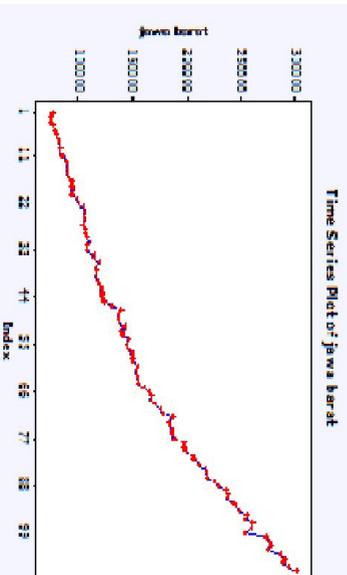
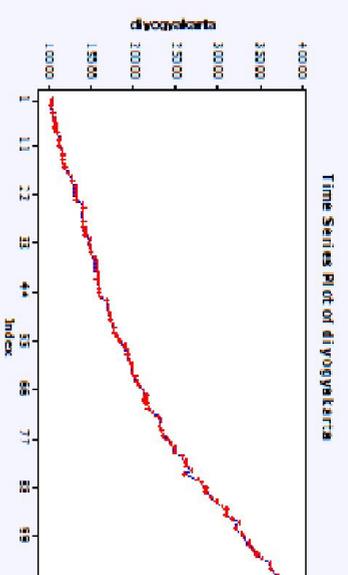
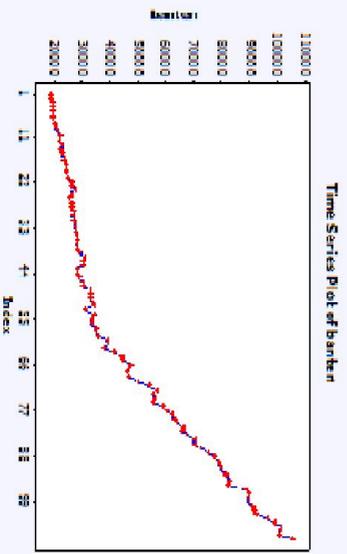
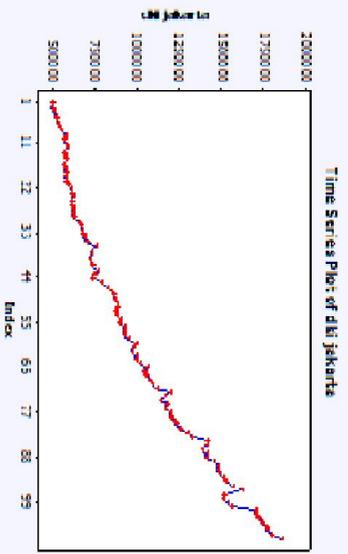
## 9. Dana Pihak Ketiga Tahun 2013

Provinsi	Jan	Feb	Maret	April	Mei	Juni	Juli	Agst	Sept	Oktr	Nov	Des
Jawa Barat	259753	259343	252235	272114	275893	277335	274398	285343	290483	288358	292379	300571
Banten	89397	89148	89124	90875	91753	92441	96176	99023	101458	100998	100665	105532
Kakarta	1519163	1512813	1536099	1559401	1707249	1707728	1723701	1749024	1769512	1776536	1796490	860446
Jogyakarta	32045	32739	33085	33597	33738	34159	34597	35237	36087	36188	36678	37073
Jawa Tengah	144307	144314	145362	148214	149491	152013	153896	156054	162827	160723	154096	167398
Jawa Timur	290230	290282	293347	296088	298121	300502	302910	307455	317370	315772	324152	335308
Bengkulu	7305	7458	7569	7895	7831	8070	8165	8115	8379	8229	8222	7580
Rambi	17923	18336	18375	18741	18458	19155	18579	18367	19521	19478	19378	19415
Aceh	21259	21307	21788	24085	23581	23736	22595	22328	26141	25052	25211	24253
Sumatra Utara												
Utarra	138709	137114	137289	136560	135515	139131	137291	140796	147985	146517	147898	155204
Sumatra Barat	23925	25793	26405	26589	26335	25906	25914	26045	26780	26977	27070	26281
Riau	50836	53216	52754	54797	53913	55990	55243	54145	56878	56425	55581	55524
Sumatra Selatan	52768	53276	53011	53803	52443	53982	52543	51422	53833	53225	52833	53588
Bangka Belitung	12387	12774	12612	12739	12429	12563	12254	12283	12693	12424	12310	11828
Lampung	22627	22835	23201	23722	23951	24349	24137	24363	25525	25892	26116	25542
Kalimantan Selatan	34020	34256	34489	35290	35254	35795	35255	34480	36299	36011	36117	36447

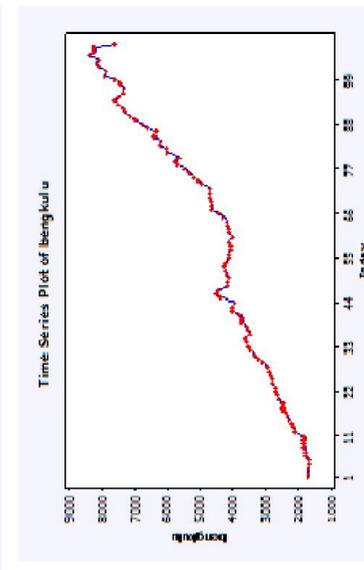
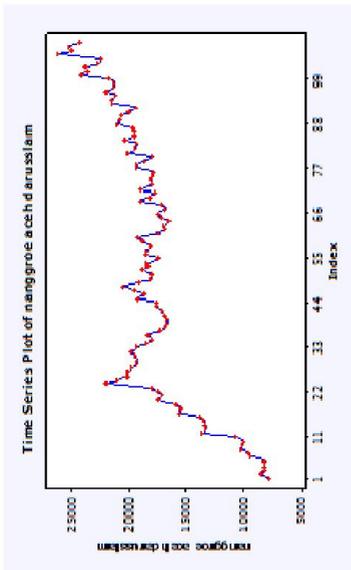
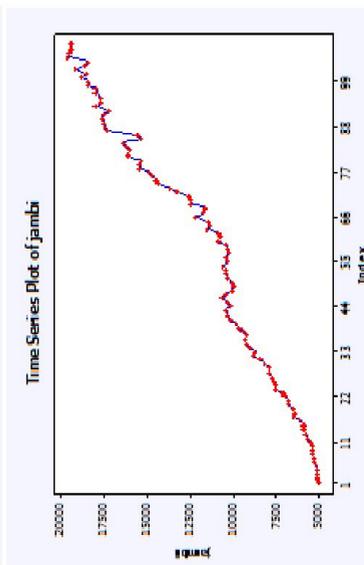
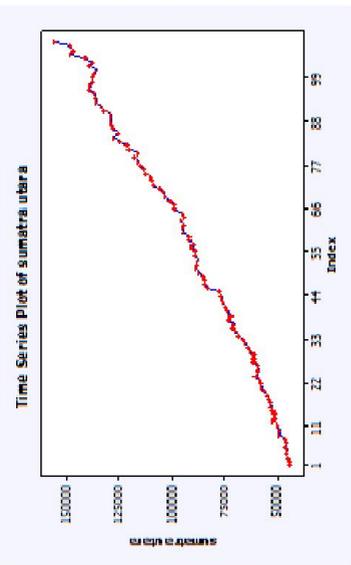
Lampiran 1. Lanjutan

Provinsi	Jan	Feb	Maret	April	Mai	Juni	Juli	Agst	Sept	Okt	Nov	Des
Kalimantan Barat	31592	31834	32407	32849	33243	33509	33921	33578	34720	34831	35455	35273
Kalimantan Timur	78591	82410	82037	84299	82554	85228	84958	83557	88869	86432	85001	85407
Kalimantan Tengah	14742	15015	15497	15341	15343	15247	15154	15908	16738	16534	16245	16023
Sulawesi Tengah	10047	10115	10442	10745	10837	10924	10927	10704	11364	11273	11280	11330
Sulawesi Selatan	53453	53157	53884	54292	54212	55220	55636	56464	58524	58841	59753	61163
Sulawesi Utara	16233	16122	16402	16803	16837	16963	17152	17039	17625	17537	17720	17481
Gorontalo	3021	3064	3152	3233	3196	3178	3186	3112	3361	3340	3331	3100
Sulawesi Tenggara	10104	10235	10578	10906	10985	10854	10738	10780	11284	11502	11665	11033
NTB	12825	12519	12749	13743	13502	13378	13388	13511	14318	14442	14648	14445
Bali	55298	55297	55982	56344	56545	57840	59042	60237	62259	62329	63615	64234
NTT	15219	15179	15351	15594	15602	15836	15906	15546	15923	16252	16694	16402
Maluku	8536	8410	8504	8651	8720	8724	8790	8713	9061	9076	9359	9379
Papua	26893	26702	26162	28150	28247	28047	30104	29791	30065	29587	30255	29424
Maluku Utara	4757	4779	4793	4923	4847	4744	4729	4757	4523	4823	5012	4831

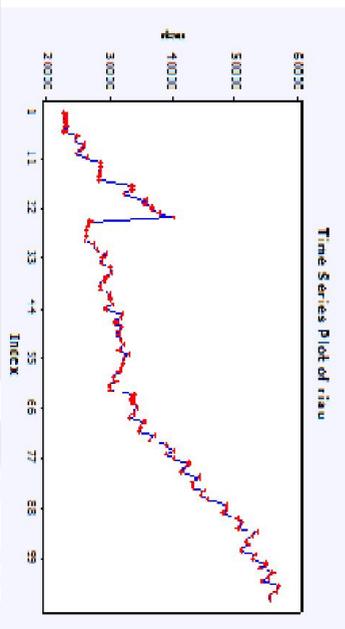
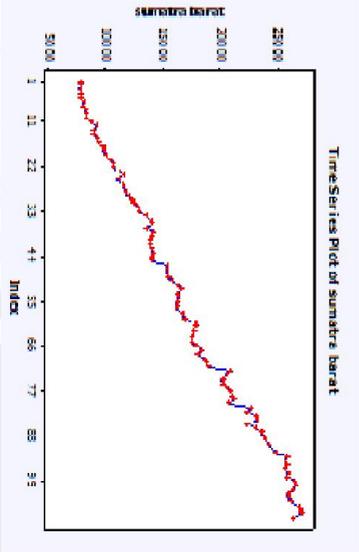
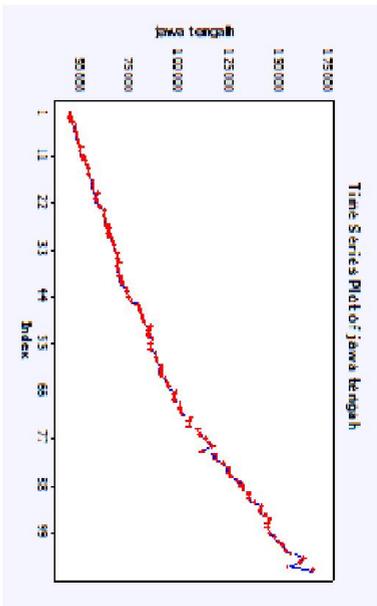
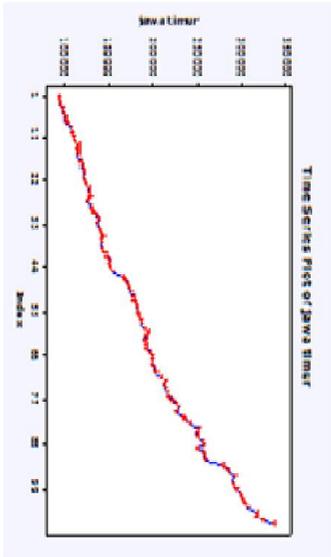
## Lampiran 2. Plot data Data Pihak Ketiga 30 provinsi



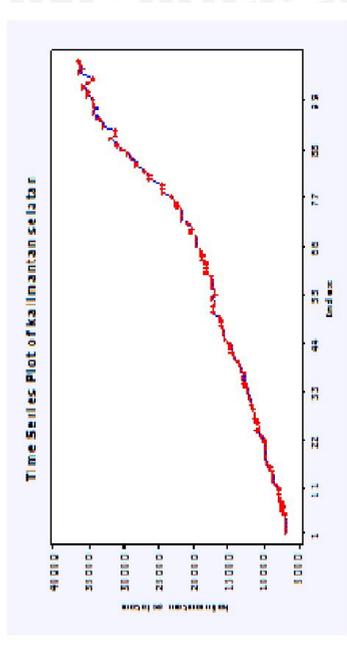
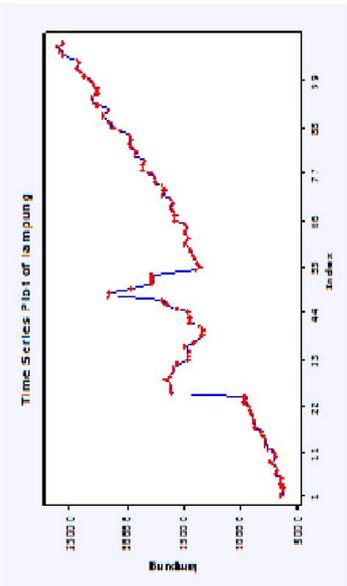
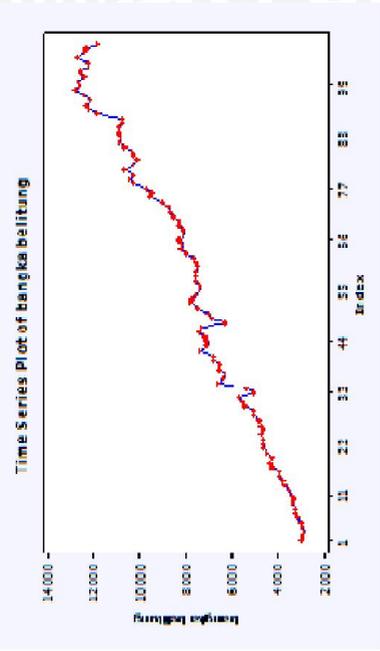
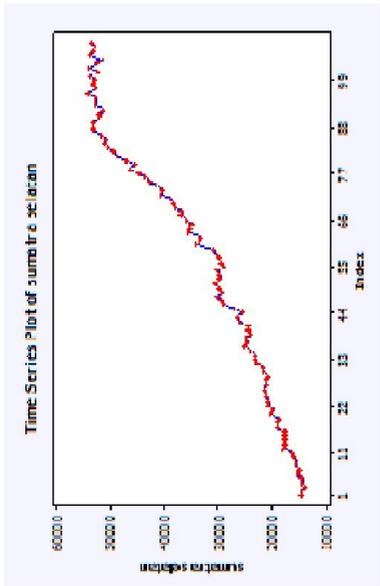
Lampiran 2. Lanjutan



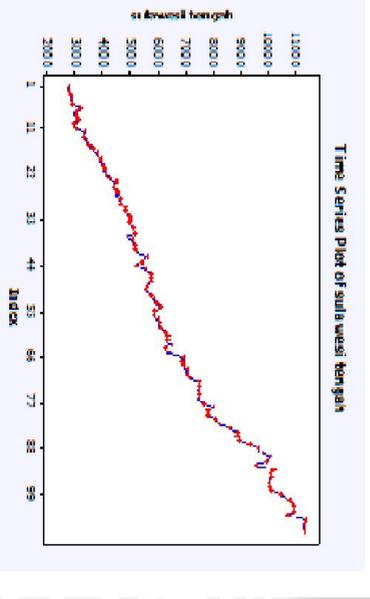
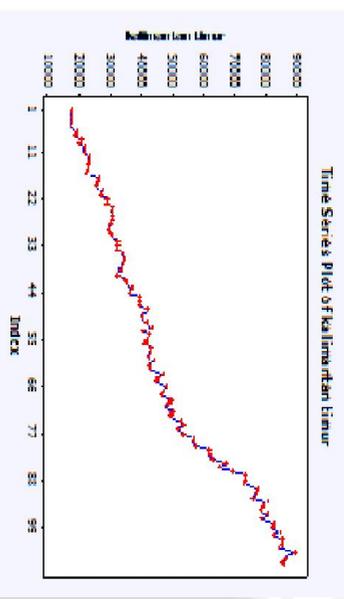
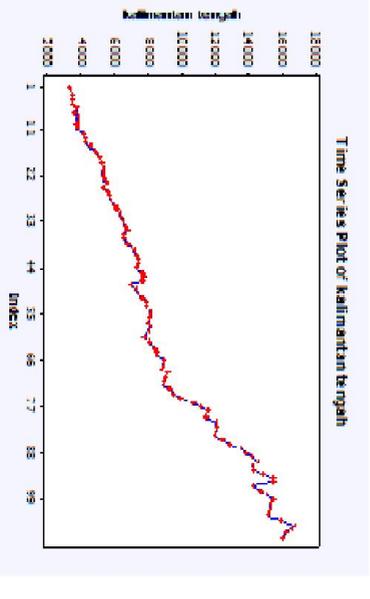
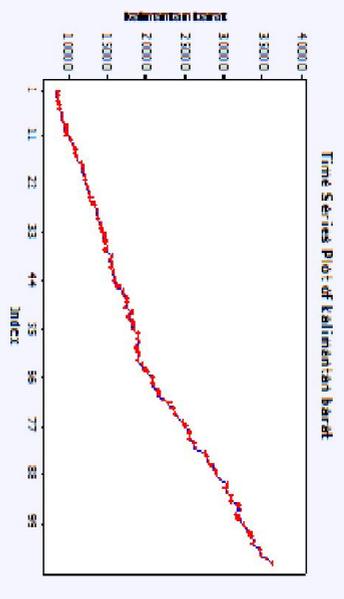
Lampiran 2: lanjut

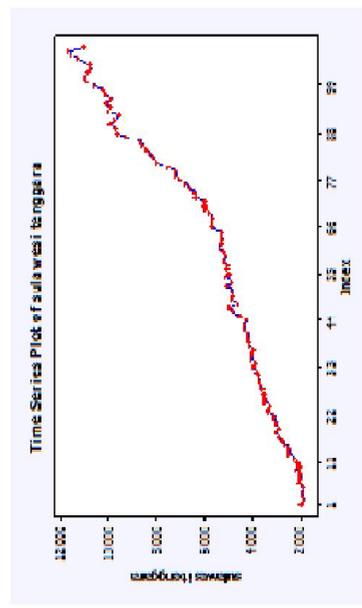
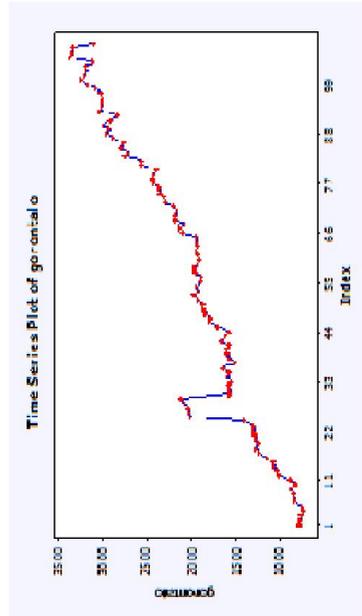
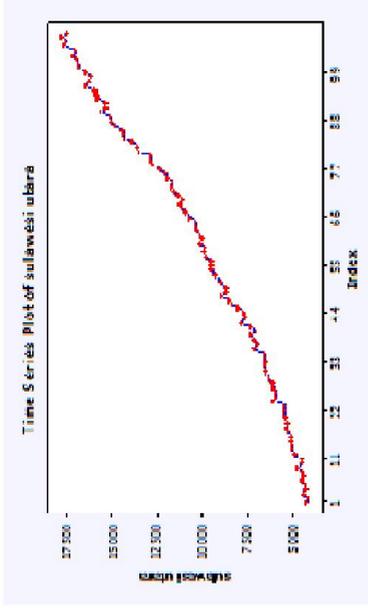
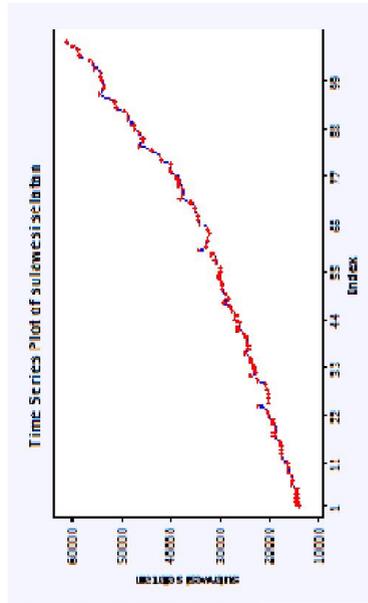


Lampiran 2. Lanjutan

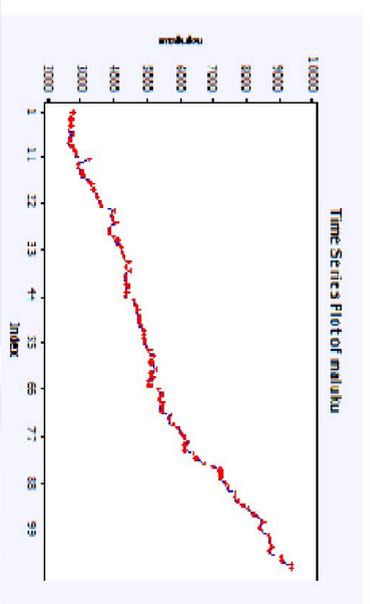
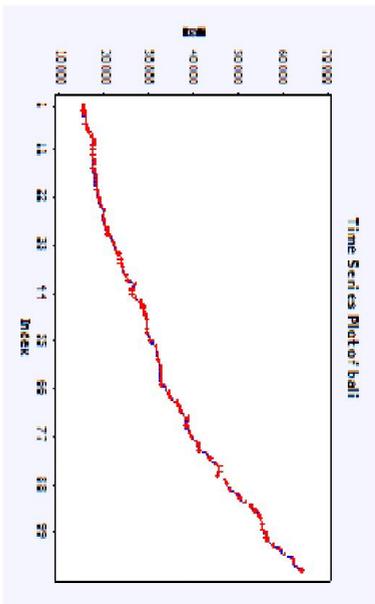
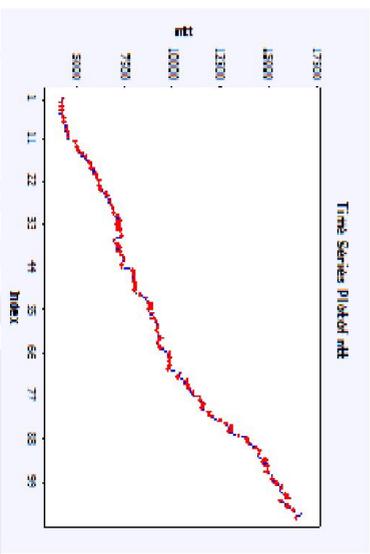
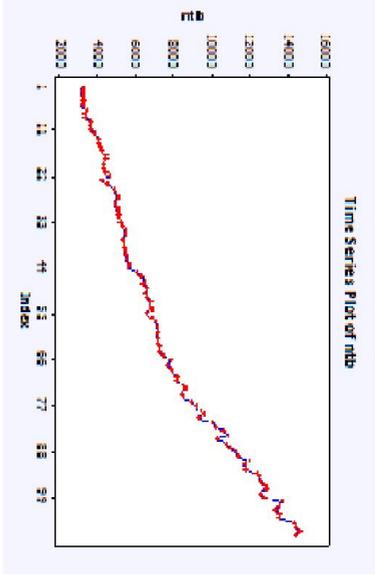


## Lampiran 2. Lanjutan

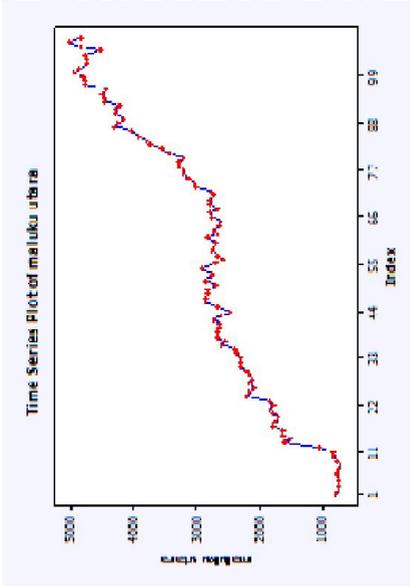
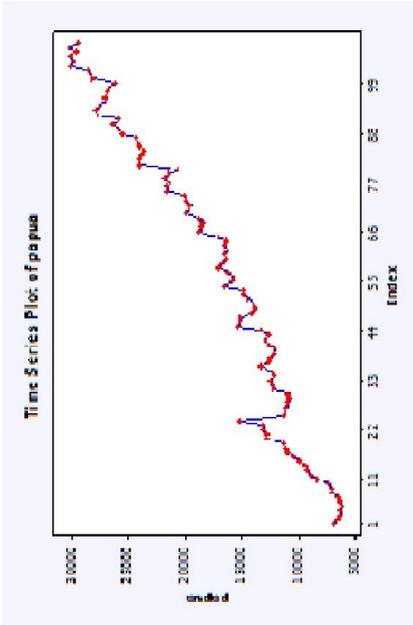




## Lampiran 2. Lanjutan



Lampiran 2. Lanjutan



UNIVERSITAS BRAWIJAYA

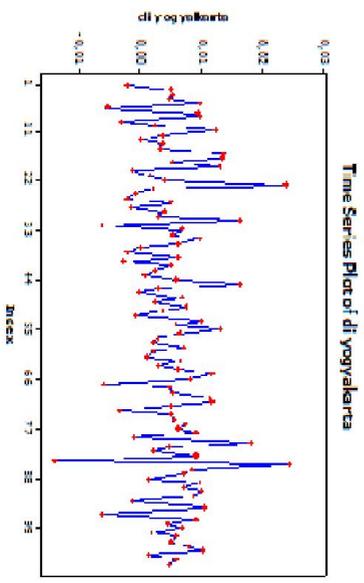
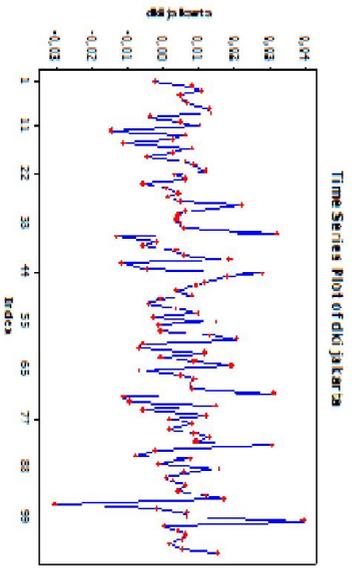
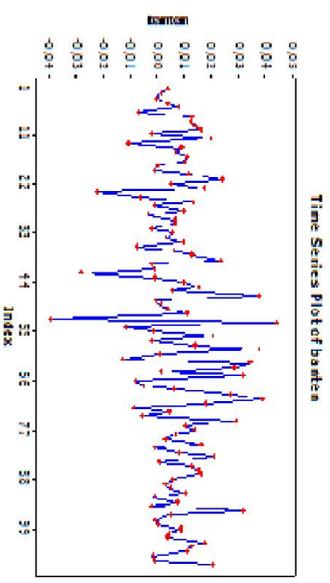
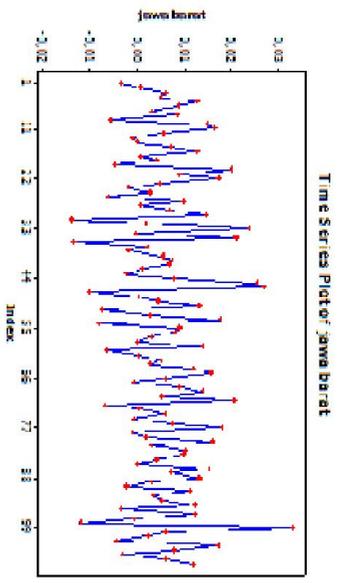
Lampiran 3. Nilai lambda data Dana Pihak Ketiga (DPK) 30 Provinsi

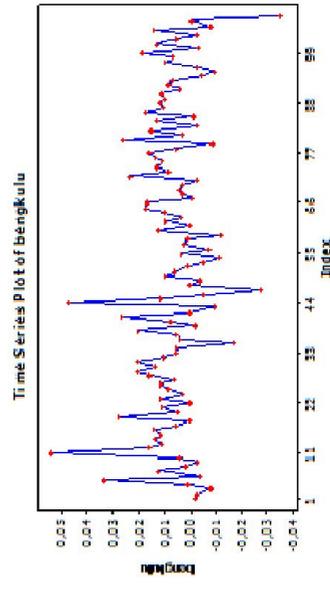
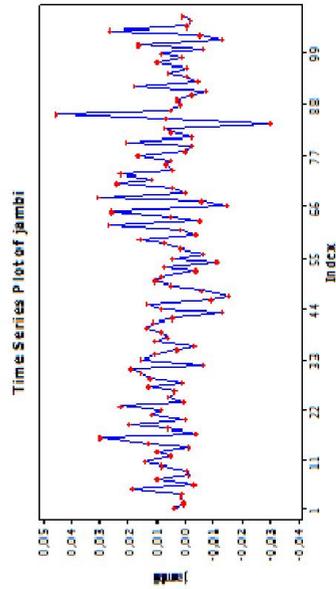
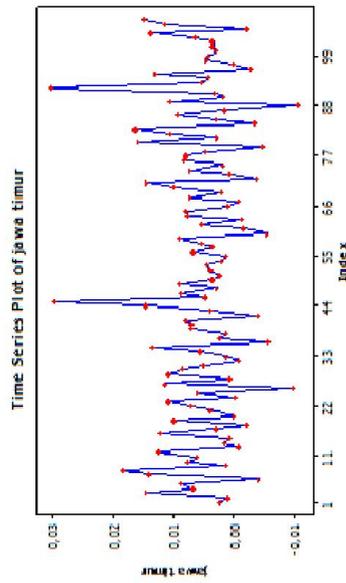
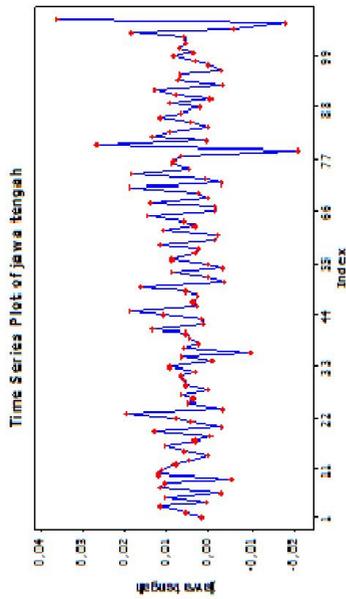
No	Provinsi	Nilai Lambda
1	Jawa Barat	0
2	Banten	0
3	DKI Jakarta	0
4	DI Yogyakarta	0
5	Jawa Tengah	-0,5
6	Jawa Timur	0
7	Bengkulu	0
8	Jambi	0
9	Nanggroe Aceh	0,5
10	Sumatra Utara	0
11	Sumatra Barat	0
12	Riau	0
13	Sumatra Selatan	0
14	Bangka Belitung	0,5
15	Lampung	0,5
16	Kalimantan Selatan	0
17	Kalimantan Barat	0
18	Kalimantan Timur	0
19	Kalimantan Tengah	0
20	Sulawesi Tengah	0,5
21	Sulawesi Selatan	0,5
22	Sulawesi Utara	0,5
23	Gorontalo	0,5
24	Sulawesi Tenggara	0,5
25	Nusa Tenggara Barat	0
26	Bali	0
27	Nusa Tenggara Timur	0
28	Maluku	0,5
29	Papua	0,5
30	Maluku utara	0,5

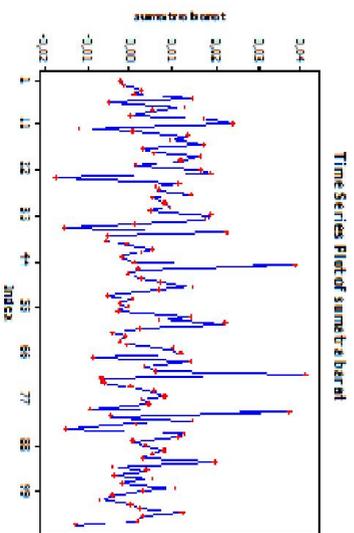
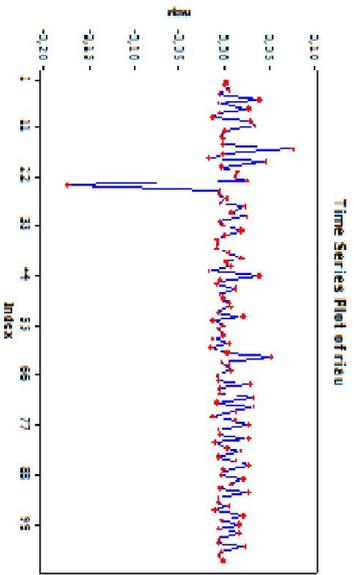
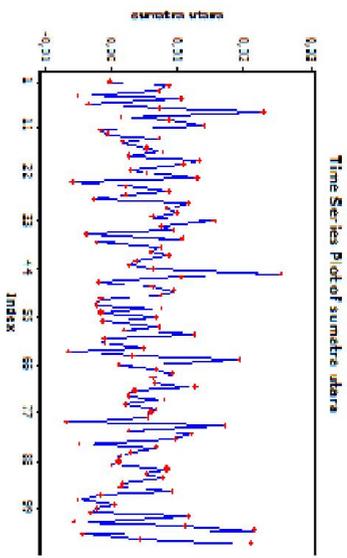
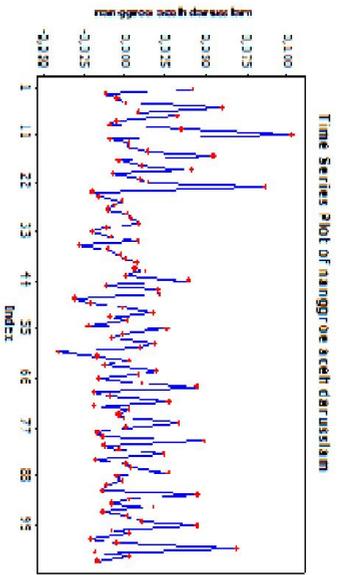
Lampiran 4. Hasil Uji Augmented Dicky Fuller data DPK 30 Provinsi

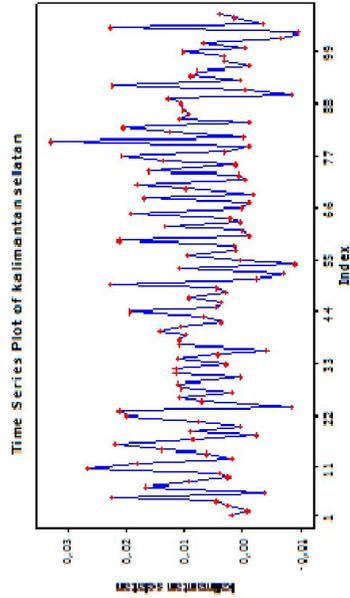
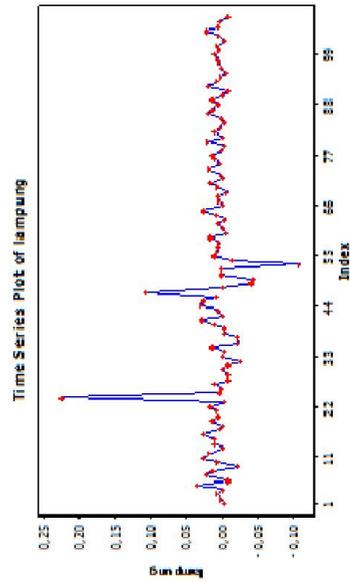
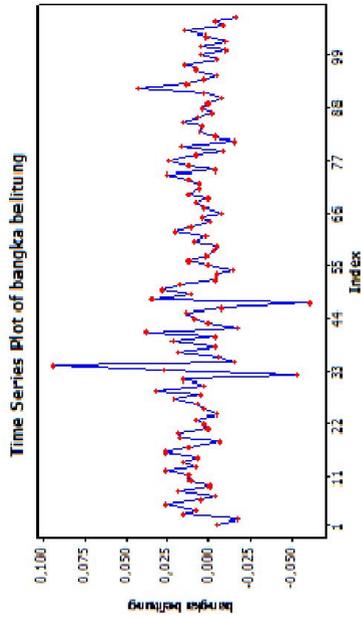
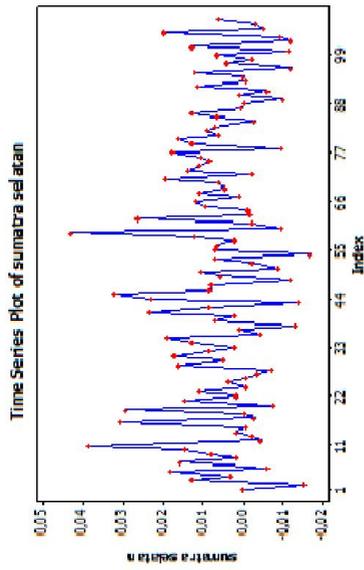
Provinsi	ADF-test	P-value	Keterangan
Jawa Barat	1,1695	0,99	Tidak Stasioner
Banten	0,513	0,99	Tidak Stasioner
DKI Jakarta	0,0307	0,99	Tidak Stasioner
DI Yogyakarta	1,7008	0,99	Tidak Stasioner
Jawa Tengah	0,2475	0,99	Tidak Stasioner
Jawa Timur	0,5494	0,99	Tidak Stasioner
Bengkulu	-1,8964	0,6196	Tidak Stasioner
Jambi	-1,5691	0,7554	Tidak Stasioner
Nanggroe Aceh	-2,6238	0,3179	Tidak Stasioner
Sumatra Utara	-2,2532	0,2716	Tidak Stasioner
Sumatra Barat	-2,196	0,4953	Tidak Stasioner
Riau	-1,6634	0,7163	Tidak Stasioner
Sumatra Selatan	-1,681	0,709	Tidak Stasioner
Bangka Belitung	-2,308	0,4468	Tidak Stasioner
Lampung	-2,9943	0,1642	Tidak Stasioner
Kalimantan Selatan	-1,2418	0,8911	Tidak Stasioner
Kalimantan Barat	0,0393	0,99	Tidak Stasioner
Kalimantan Timur	-1,3586	0,8427	Tidak Stasioner
Kalimantan Tengah	-1,1219	0,9154	Tidak Stasioner
Sulawesi Tengah	-0,6243	0,9747	Tidak Stasioner
Sulawesi Selatan	0,251	0,99	Tidak Stasioner
Sulawesi Utara	-2,0256	0,566	Tidak Stasioner
Gorontalo	-2,9853	0,1679	Tidak Stasioner
Sulawesi Tenggara	-0,8502	0,9545	Tidak Stasioner
Nusa Tenggara Barat	0,4454	0,99	Tidak Stasioner
Bali	0,3931	0,99	Tidak Stasioner
Nusa Tenggara Timur	-0,6189	0,9751	Tidak Stasioner
Maluku	0,1036	0,99	Tidak Stasioner
Papua	-2,1241	0,5252	Tidak Stasioner
Maluku utara	-1,9423	0,6006	Tidak Stasioner

**Lampiran 5. Plot data deret waktu setelah dilakukan transformasi**

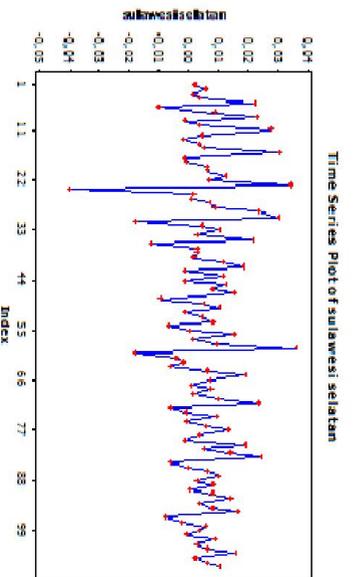
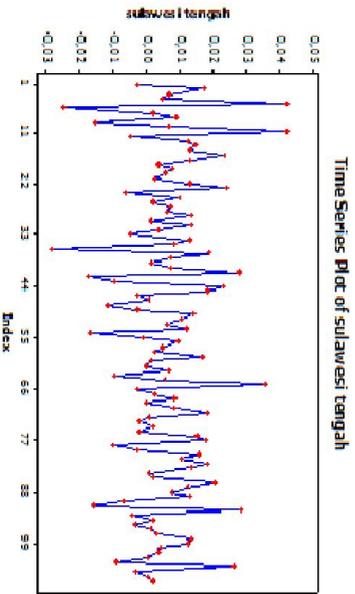
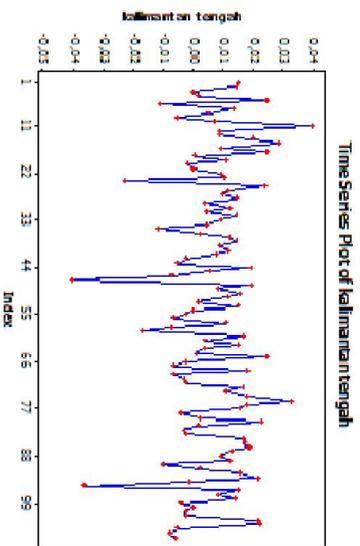
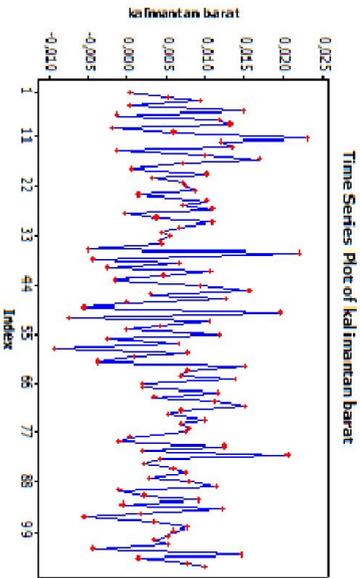


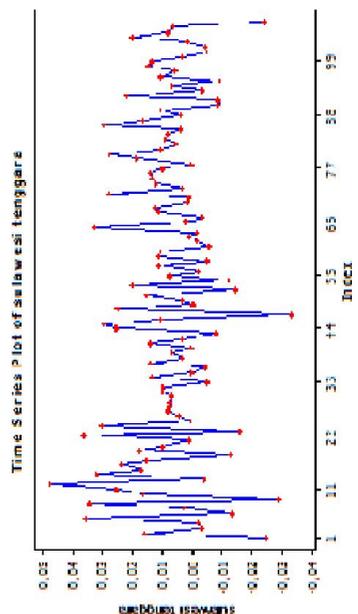
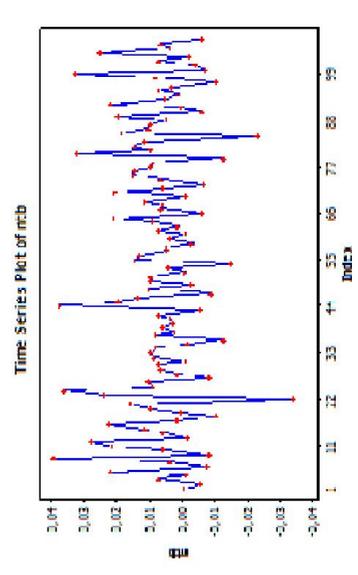
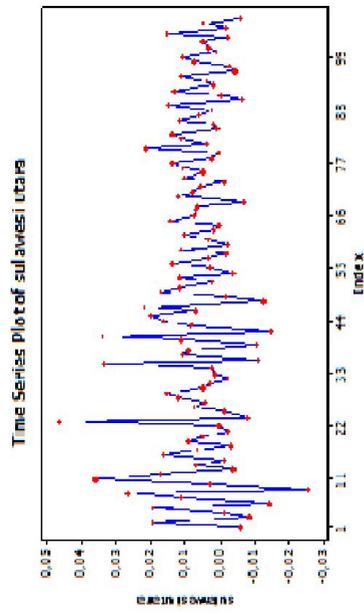
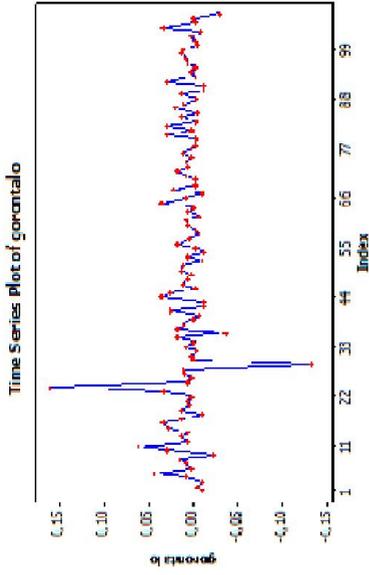




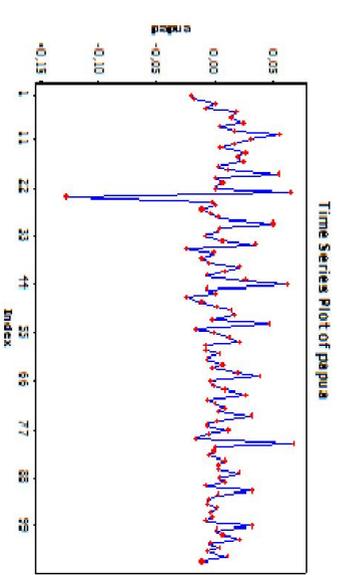
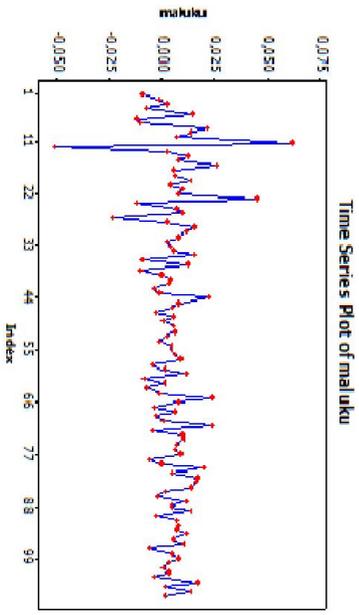
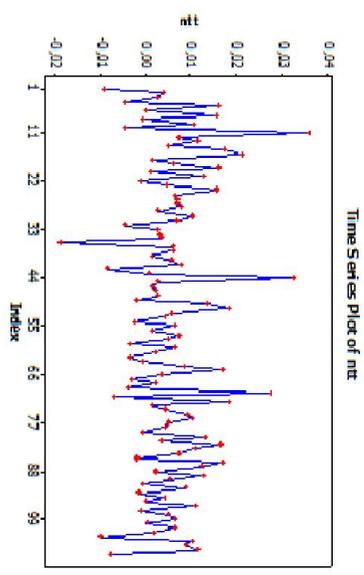
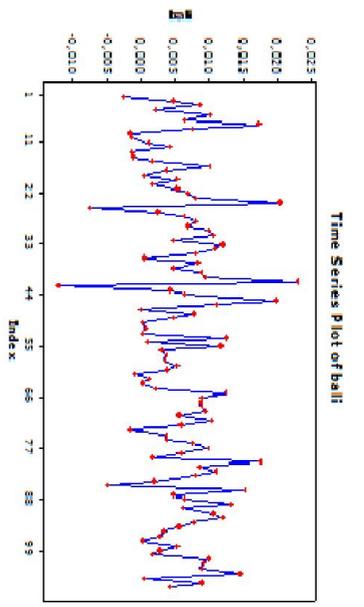


Lampiran 5. Lanjutan

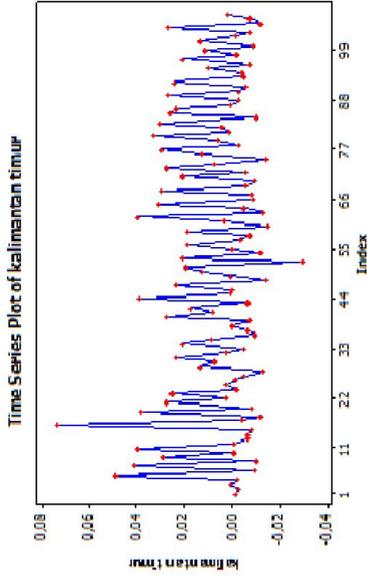
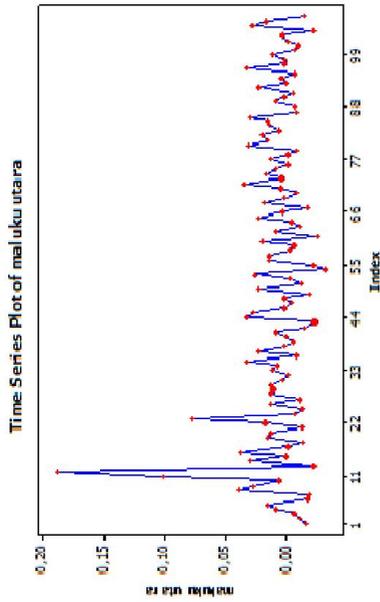




Lampiran 5. Lanjutan



Lampiran 5. Lanjutan



S BRAWIJAYA

Lampiran 6. Hasil Uji Augmented Dicky Fuller setelah transformasi

Provinsi	ADF-test	P-value	Keterangan
Jawa Barat	-6,1833	0,01	Stasioner
Banten	-6,4239	0,01	Stasioner
DKI Jakarta	-6,198	0,01	Stasioner
DI Yogyakarta	-5,7476	0,01	Stasioner
Jawa Tengah	-5,5088	0,01	Stasioner
Jawa Timur	-4,936	0,01	Stasioner
Bengkulu	-3,9531	0,01	Stasioner
Jambi	-4,6239	0,01	Stasioner
Nanggroe Aceh	-4,0025	0,0117	Stasioner
Sumatra Utara	-3,9303	0,0152	Stasioner
Sumatra Barat	-5,44	0,01	Stasioner
Riau	-5,2607	0,01	Stasioner
Sumatra Selatan	-4,5706	0,01	Stasioner
Bangka Belitung	-5,2585	0,01	Stasioner
Lampung	-4,2693	0,01	Stasioner
Kalimantan Selatan	-4,4779	0,01	Stasioner
Kalimantan Barat	-4,03	0,01037	Stasioner
Kalimantan Timur	-4,852	0,01	Stasioner
Kalimantan Tengah	-4,4585	0,01	Stasioner
Sulawesi Tengah	-6,8908	0,01	Stasioner
Sulawesi Selatan	-7,33	0,01	Stasioner
Sulawesi Utara	-6,6347	0,01	Stasioner
Gorontalo	-5,7771	0,01	Stasioner
Sulawesi Tenggara	-3,7543	0,02373	Stasioner
Nusa Tenggara Barat	-5,0395	0,01	Stasioner
Bali	-4,7649	0,01	Stasioner
Nusa Tenggara Timur	-3,938	0,01483	Stasioner
Maluku	-5,872	0,01	Stasioner
Papua	-4,6603	0,01	Stasioner
Maluku utara	-4,5767	0,01	Stasioner

Lampiran 7. Jarak Penggabungan Masing-Masing *Stage* Menggunakan Jarak *Autocorrelation-based* Pada Analisis *Cluster* Hirarki

<i>Stage</i>	Koefisien	Selisih
1	0,003891551	
2	0,004167402	0,00027585
3	0,005073837	0,00090644
4	0,005153294	0,00007946
5	0,005376443	0,00022315
6	0,005999193	0,00062275
7	0,006170041	0,00017085
8	0,006482326	0,00031229
9	0,006577497	0,00009517
10	0,007557435	0,00097994
11	0,008124989	0,00056755
12	0,008405897	0,00028091
13	0,008522049	0,00011615
14	0,008907972	0,00038592
15	0,009039868	0,00013190
16	0,010351771	0,00131190
17	0,010620765	0,00026899
18	0,011232062	0,00061130
19	0,011405994	0,00017393
20	0,011751278	0,00034528
21	0,012775562	0,00102428
22	0,012845398	0,00006984
23	0,013726858	0,00088146
23	0,014719545	0,00099269
25	0,015001140	0,00028159
26	0,015517912	0,00051677
27	0,016810662	0,00129275
28	0,020239236	0,00342857
29	0,033622895	0,01338366

Lampiran 8. Jarak Penggabungan Masing-Masing *Stage* Menggunakan Jarak *Complexity Invariant* Pada Analisis *Cluster* Hirarki

<i>Stage</i>	Koefisien	Selisih
1	0,1469727	
2	0,1613417	0,014369
3	0,1775879	0,016246
4	0,1861151	0,008527
5	0,2022440	0,016129
6	0,2141980	0,011954
7	0,2368330	0,022635
8	0,2568479	0,020015
9	0,2606248	0,003777
10	0,2627504	0,002126
11	0,2740323	0,011282
12	0,2850213	0,010989
13	0,2879498	0,002929
14	0,2950593	0,00711
15	0,3268698	0,031811
16	0,3439822	0,017112
17	0,3527820	0,0088
18	0,3706087	0,017827
19	0,4249965	0,054388
20	0,4877706	0,062774
21	0,4907031	0,002933
22	0,5161877	0,025485
23	0,5412064	0,025019
23	0,6011291	0,059923
25	0,6680981	0,066969
26	0,6743254	0,006227
27	0,7594917	0,085166
28	0,9496289	0,190137
29	1,3967529	0,447124

Lampiran 9. Jarak Penggabungan Masing-Masing *Stage* Pada Analisis *Cluster* Hirarki Menggunakan Jarak *Periodogram-based*

<i>Stage</i>	Koefisien	Selisih
1	0,1104195	
2	0,1171903	0,006771
3	0,1193568	0,002167
4	0,1207688	0,001412
5	0,1292161	0,008447
6	0,1321407	0,002925
7	0,1323295	0,000189
8	0,1377234	0,005394
9	0,1394176	0,001694
10	0,1412318	0,001814
11	0,1421111	0,000879
12	0,1492116	0,007101
13	0,1513726	0,002161
14	0,1536485	0,002276
15	0,1563889	0,00274
16	0,1680316	0,011643
17	0,1704199	0,002388
18	0,1752339	0,004814
19	0,1755045	0,000271
20	0,1770031	0,001499
21	0,1774490	0,000446
22	0,1808381	0,003389
23	0,1817243	0,000886
23	0,1842655	0,002541
25	0,1905758	0,00631
26	0,1916630	0,001087
27	0,2048426	0,01318
28	0,2275783	0,022736
29	0,2737790	<b>0,046201</b>

Lampiran 10. Jarak Penggabungan Masing-Masing *Stage* Pada Analisis *Cluster* Hirarki Menggunakan Jarak *Dynamic Time Warping*

<i>Stage</i>	Koefisien	Selisih
1	1,158294	
2	1,283253	0,124959
3	1,347616	0,064363
4	1,362696	0,01508
5	1,481790	0,119094
6	1,599534	0,117744
7	1,672434	0,0729
8	1,716208	0,043774
9	1,918214	0,202006
10	1,966133	0,047919
11	2,038081	0,071948
12	2,089182	0,051101
13	2,108507	0,019325
14	2,138528	0,030021
15	2,218774	0,080246
16	2,240075	0,021301
17	2,292977	0,052902
18	2,307071	0,014094
19	2,431913	0,124842
20	2,492602	0,060689
21	2,753697	0,261095
22	2,794080	0,040383
23	2,923698	0,129618
23	3,309645	0,385947
25	3,393458	0,083813
26	3,553484	0,160026
27	3,612233	0,058749
28	3,951741	0,339508
29	4,366541	0,4148

Lampiran 11. Nilai Koefisien *Silhouette* Pada Tiap jumlah kelompok yang Mungkin Pada Analisis *Cluster* Menggunakan Jarak *Dynamic Time Warping* dengan Algoritma *K-Means*

Jumlah Cluster	Koefisien Silhouette
2	0,2574722
3	0,2133704
4	-0,004318871
5	-0,003560235
6	0,00455472
7	0,006781414
8	0,02520842
9	0,03328317
10	0,03608278
11	0,03901368
12	0,03523255
13	0,04064082
14	0,03751245
15	0,04732568
16	0,0504485
17	0,04892587
18	0,01316082
19	0,01802886
20	0,02541946
21	0,01958558
22	0,02492321
23	0,02443736
24	0,02260292
25	0,00294393
26	0,002177394
27	0,005544208
28	0,01611498
29	-0,005498461

## Lampiran 12. Syntax Analisis Cluster Data Deret Waktu

<code>trans=log(data[r,c])-log(data1[r,c])</code>	Melakukan transformasi log diff
<code>library(dtw)</code>	Memanggil package DTW
<code>jarak=dist(trans, method="DTW")</code>	Perhitungan jarak DTW
<code>names(jarak)=colnames(data)</code>	Memasangkan jarak menjadi kolom dari tiap objek pada data
<code>cluster=hclust(jarak,method="average")</code>	Melakukan pengelompokkan menggunakan metode hirarki <i>average linkage</i>
<code>cluster\$height</code>	Memunculkan koefisien penggabungan tiap <i>stage</i>
<code>cor(cophenetic(cluster), jarak)</code>	Menghitung koefisien korelasi <i>Cophenetic</i>
<code>library(TSclust)</code>	Memanggil package <i>Time Series Clustering</i>
<code>jarak=diss(t(trans),"PER")</code>	Menghitung jarak <i>periodogram-based</i>
<code>jarak=diss(t(trans),"CID")</code>	Menghitung <i>Complexity Invariant Distance</i>
<code>jarak=diss(t(trans),"ACF")</code>	Menghitung jarak <i>autocorrelation-based</i>
<code>Kmeans(jarak, k)</code>	Melakukan pengelompokkan dengan algoritma <i>k-means</i> sebanyak <i>k</i> kelompok
<code>Library(fpc)</code>	Memanggil package <i>flexible procedures for clustering</i>
<code>Cluster.stats(jarak,kmeans\$cluster)</code>	Memunculkan <i>average silhouette</i> dari cluster <i>k-means</i>
<code>Cluster.stat\$avg.silwidth</code>	Memunculkan dan menghitung nilai Koefisien <i>silhouette</i>