

## PENGELOMPOKAN NEGARA DI DUNIA BERDASARKAN DATA RUNTUN WAKTU REALISASI PENANAMAN MODAL ASING DI INDONESIA MENGUNAKAN ANALISIS CLUSTER

AFRIMAYANI, HAZMIRA YOZZA, DODI DEVIANTO

*Program Studi S1 Matematika,  
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,  
Kampus UNAND Limau Manis Padang, Indonesia.  
email : afrimayani1997@gmail.com*

Diterima 22 Juni 2019    Direvisi 6 Juli 2019    Dipublikasikan 4 Agustus 2019

**Abstrak.** Investasi sangat dibutuhkan di Indonesia. Perekonomian di Indonesia masih tertinggal dari negara-negara yang lain. Oleh karena itu, pemerintah perlu mencari sumber pembiayaan pembangunan ekonomi untuk mendorong pertumbuhan investasi, bukan hanya investasi dari dalam negeri namun juga investasi asing. Untuk melihat pola besarnya investasi asing di Indonesia, perlu dilakukan pengelompokan negara-negara. Pengelompokan tersebut diharapkan bisa membantu pemerintah dalam pengambilan kebijakan terkait investasi asing sesuai dengan negaranya. Analisis *cluster* merupakan suatu teknik analisis statistik dengan tujuan untuk memilah objek ke dalam beberapa *cluster* berdasarkan kesamaan-kesamaan objek atas dasar berbagai karakteristik. Penelitian ini bertujuan untuk membentuk *cluster* negara-negara di dunia berdasarkan data runtun waktu realisasi investasinya di Indonesia tahun 2000-2017. Teknik pengelompokan yang digunakan adalah analisis berhierarki dengan jarak *euclidean*. Hasil penelitian ini diperoleh sebanyak 2 *cluster* sebagai *cluster* optimum. Pada *cluster* 1 besarnya realisasi penanaman modal asing sepanjang tahun 2000-2017 cenderung hampir sama besar. Pada *cluster* 2 besarnya realisasi penanaman modal asing sepanjang tahun 2000-2017 lebih besar dibandingkan dengan *cluster* 1 dan mengalami perubahan yang signifikan.

*Kata Kunci:* Investasi Asing, Analisis *Cluster* Runtun Waktu, Jarak *Euclidean*

### 1. Pendahuluan

Penanaman modal (investasi) sangat dibutuhkan di setiap negara, demikian juga halnya dengan Indonesia. Perekonomian di Indonesia masih tertinggal dari negara-negara yang lain. Oleh karena itu, pemerintah perlu mencari sumber-sumber pembiayaan pembangunan ekonomi untuk mendorong pertumbuhan investasi. Indonesia sangat membutuhkan investasi, bukan hanya investasi dari dalam negeri namun juga investasi asing. Salah satu alasan utama mengapa Indonesia membutuhkan investor asing adalah daya konsumsi masyarakat serta tren ekspor yang masih rendah.

Pemerintah perlu mempertahankan strategi dalam menarik investor asing yang mendorong mereka untuk berinvestasi di sektor yang belum mampu diolah oleh pihak dalam negeri. Untuk melihat pola besarnya investasi asing di Indonesia, perlu

dilakukan pengelompokan negara-negara. Pengelompokan tersebut diharapkan bisa membantu pemerintah dalam pengambilan kebijakan terkait investasi asing sesuai dengan negaranya. Dalam menentukan strategi untuk mengelompokkan negara asing, diperlukan analisis untuk mengelompokkan negara yang tepat. Dalam konsep statistik, suatu analisis yang digunakan untuk mengelompokkan objek adalah analisis *cluster*.

Analisis *cluster* merupakan suatu teknik analisis statistik dengan tujuan untuk memilah objek ke dalam beberapa *cluster* berdasarkan kesamaan-kesamaan objek atas dasar berbagai karakteristik [7]. Ukuran jarak yang digunakan adalah ukuran jarak *euclidean*. Selain itu, proses pengelompokan ini juga dibutuhkan metode perbaikan jarak yaitu metode pautan lengkap (*complete linkage*). Pada tulisan ini akan dilakukan pengelompokan negara-negara di dunia berdasarkan besar investasi yang ditanamkan di Indonesia dalam beberapa tahun menggunakan analisis *cluster*.

## 2. Landasan Teori

### 2.1. Investasi

Investasi merupakan bentuk penundaan konsumsi masa sekarang untuk memperoleh konsumsi di masa yang akan datang, dimana di dalamnya terkandung unsur risiko ketidakpastian sehingga dibutuhkan kompensasi atas penundaan tersebut [5]. Investasi ada dua macam, investasi dalam negeri dan investasi luar negeri (asing). *Risk* dan *return* merupakan prinsip utama dari investasi. *Return* adalah laba dari investasi sedangkan *risk* adalah kemungkinan mengalami kerugian dalam investasi. Pada umumnya hubungan antara *return* dan *risk* adalah linear. Umumnya, risiko yang lebih rendah diambil, keuntungan yang diperoleh lebih rendah dan sebaliknya. Prinsip-prinsip ini penting bagi investor sebelum membuat kesepakatan investasi [2]. Menurut UU No. 1 Th. 1967, yang dimaksud dengan "*Penanaman Modal Asing adalah penanaman modal asing secara langsung yang dilakukan menurut atau berdasarkan ketentuan-ketentuan undang-undang ini dan yang digunakan untuk menjalankan perusahaan di Indonesia, dalam arti bahwa pemilik modal secara langsung menanggung risiko dari penanaman modal tersebut*".

### 2.2. Analisis Cluster

*Clustering* adalah proses pengelompokan objek ke dalam kelas-kelas yang memiliki kesamaan. *Clustering* didasarkan pada kesamaan (*similarity*) atau ketaksamaan (*dissimilarity*) antar objek-objek. Satu-satunya cara untuk mengukur ketaksamaan adalah dengan suatu ukuran jarak (*distance measure*). Hasil dari proses *clustering* adalah *cluster-cluster* yang terdiri dari sekumpulan objek dengan kesamaan tertentu dalam satu *cluster* dan berbeda dengan objek yang berada pada *cluster* lain [7].

Ada dua metode *cluster* yaitu metode hierarki dan metode nonhierarki [6]. Metode hierarki diklasifikasikan ke dalam dua jenis yaitu *Agglomerative* (penyatuan) dan *Divisive* (pembagian). Perbedaan dari kedua metode ini bergantung pada komposisi hirarkis yang dilakukan, apakah dari bentuk *bottom-up* atau *top-down*. Metode yang digunakan pada penelitian ini adalah metode hierarki jenis *agglomerative*. Dalam metode hierarki, jarak masing-masing objek dihitung terhadap

setiap objek yang lain. Selanjutnya, pasangan objek yang jaraknya terdekat digabung [7]. Tahap berikutnya dilakukan metode perbaikan jarak. Salah satu metode perbaikan jarak yang digunakan adalah metode pautan lengkap (*complete linkage*).

Metode pautan lengkap (*complete linkage*) didasarkan pada jarak maksimum. Jarak antara satu *cluster* dengan *cluster* yang lain diukur berdasarkan objek yang mempunyai jarak terjauh. Misalnya *cluster*  $U$  dan  $V$ , untuk mendapatkan *cluster*  $(UV)$ . Jarak antara *cluster*  $(UV)$  dan *cluster* lain  $W$ , dapat ditulis ke dalam rumus sebagai berikut [3]:

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\}, \quad (2.1)$$

dimana  $d_{UW}$  dan  $d_{VW}$  merupakan jarak antara anggota *cluster*  $U$  dengan *cluster*  $W$  dan juga *cluster*  $V$  dengan *cluster*  $W$ .

Prosedur tersebut dilakukan sampai diperoleh satu *cluster* yang memuat semua objek dan disajikan ke dalam bentuk *dendogram*. *Dendogram* disebut juga dengan grafik pohon, yaitu suatu alat grafis untuk menyajikan hasil dari analisis *cluster* yang dilakukan oleh peneliti dalam melakukan suatu penelitian. *Dendogram* berguna untuk menunjukkan anggota *cluster* yang ada jika akan ditentukan berapa banyak *cluster* yang seharusnya dibentuk.

### 2.3. Data Runtun Waktu (*Time Series*)

Data runtun waktu merupakan himpunan hasil observasi terhadap suatu variabel yang diambil secara beruntun atau terurut berdasarkan interval waktu yang tetap. Data runtun waktu adalah data yang dikumpulkan dari waktu ke waktu. Tujuannya adalah untuk menggambarkan perkembangan suatu kegiatan dari waktu ke waktu. Rangkaian data pengamatan runtun waktu dinyatakan dengan variabel  $X_t$ , dimana  $t$  adalah indeks waktu dari urutan pengamatan [8].

### 2.4. Analisis *Cluster* Data Runtun Waktu (*Time Series*)

**Definisi 2.1.** [1] *Analisis cluster data runtun waktu adalah suatu proses untuk mempartisi data runtun waktu  $D = \{F_1, F_2, F_3, \dots, F_n\}$  menjadi  $C = \{C_1, C_2, C_3, \dots, C_k\}$  dalam suatu cara, sehingga data runtun waktu yang homogen akan dikelompokan bersamaan berdasarkan suatu ukuran kesamaan tertentu, maka  $C_i$  sebagai cluster, dimana  $D = \bigcup_{i=1}^k C_i$  dan  $C_i \cap C_j = \emptyset$  untuk setiap  $i \neq j$ .*

#### 2.4.1. Jarak *Euclidean* (*Euclidean Distance*)

Adapun rumus jarak *euclidean* untuk dua data runtun waktu  $u$  dan  $v$  adalah sebagai berikut.

$$d_{uv} = \sqrt{\sum_{t=1}^n (u_t - v_t)^2}, \quad (2.2)$$

dimana

$n$  : Panjang data  $u$  dan data  $v$ ,

$d_{uv}$  : Jarak euclidean antara data runtun waktu  $u$  dan data runtun waktu  $v$ .

### 2.5. Koefisien *Silhouette*

Misal  $n$  objek dapat dimasukkan sebanyak  $k$  *cluster*. Batasan *cluster* yang dapat dibentuk adalah  $k \leq n$ , dengan  $k = 1$  dan  $k = n$  dikecualikan, sehingga diperoleh  $2 \leq k \leq (n - 1)$ . Koefisien *silhouette* yang digunakan untuk mengetahui kualitas pengelompokan atau pengukuran ketepatan pengelompokan dapat dirumuskan sebagai berikut [4].

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (2.3)$$

dimana

- $i$  : objek yang akan diteliti,
- $a(i)$  : rata-rata jarak antar anggota dalam *cluster*,
- $b(i)$  : nilai minimum dari rata-rata jarak dari objek ke  $-i$  dengan objek yang berada di *cluster* lainnya,
- $S(i)$  : nilai koefisien *silhouette* pada objek ke  $-i$ .

### 3. Identifikasi Masalah dan Data

Negara-negara asing yang menanamkan modalnya di Indonesia berjumlah 17 negara yaitu Jerman, Swiss, Kanada, India, Belgia, Denmark, Italia, Perancis, Taiwan, Australia, Inggris, Belanda, Hongkong, USA, Jepang, Singapura dan Korea Selatan. Negara ini dipilih dengan pertimbangan kelengkapan data realisasi penanaman modal asing di Indonesia sepanjang tahun 2000-2017.

Langkah-langkah tahapan penelitian ini adalah melakukan analisis deskriptif untuk menggambarkan pergerakan investasi negara-negara asing di Indonesia, memperoleh matriks jarak *euclidean*, menggabungkan *cluster* yang memiliki jarak terdekat, melakukan perbaikan matriks jarak dengan metode *complete linkage*, mengulang langkah tersebut sebanyak  $N - 1$  kali, sehingga diperoleh satu *cluster* yang berisi semua objek, memilih jumlah *cluster* optimum yang terbentuk menggunakan kriteria koefisien *silhouette* dan melakukan interpretasi analisis *cluster* sesuai metode yang ditentukan.

### 4. Pembahasan

Data yang digunakan adalah data realisasi penanaman modal asing di Indonesia mulai dari tahun 2000 sampai tahun 2017. Kemudian dari data runtun waktu tersebut akan dilakukan pengelompokan negara-negara asing. Tahap awal yang harus dilakukan adalah menganalisis deskriptif untuk menggambarkan pergerakan investasi negara-negara asing di Indonesia. Berdasarkan data investasi asing tersebut diperoleh negara Jepang dan Singapura memiliki pola pergerakan dan besar realisasi investasi yang sangat berbeda dengan negara-negara yang lain. Sehingga dapat dikatakan bahwa kedua negara tersebut berbeda dari negara-negara lain dari segi realisasi modal yang ditanamkannya.

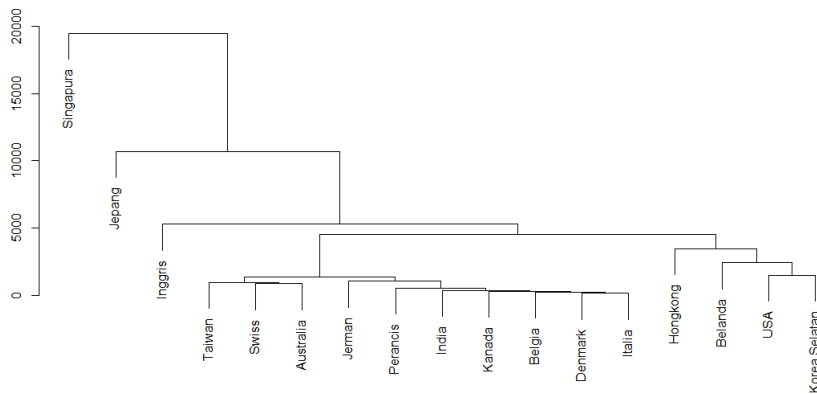
Diketahui jarak *euclidean* antara negara USA dengan negara Kanada adalah 4091.5703, sedangkan jarak *Euclidean* antara negara USA dengan negara Belgia

adalah 4173.6744. Hal ini menunjukkan bahwa negara USA memiliki karakteristik yang lebih mirip dengan negara Kanada dibandingkan dengan negara USA dengan negara Belgia.

Penafsiran terhadap ukuran jarak antara negara-negara lain sama dengan yang diperlihatkan sebelumnya, dimana semakin kecil jarak yang diperoleh antara kedua objek, maka akan semakin mirip karakteristik dari kedua objek tersebut. Proses pengelompokan diawali dengan menganggap masing-masing objek sebagai satu *cluster*. Dengan demikian akan terdapat 17 *cluster* yang masing-masingnya terdiri dari satu negara. Dari 17 *cluster* tersebut terlebih dahulu dilihat *cluster* dengan jarak terdekat. Jarak antara negara Denmark (objek d) dan negara Italia (objek f) merupakan jarak yang terdekat sebesar 148,3478, sehingga kedua negara tersebut digabung menjadi satu *cluster*. Dengan demikian, akan terdapat 16 *cluster*.

Kemudian dari penggabungan dua *cluster* di atas dilakukan perbaikan jarak dengan menggunakan metode *complete linkage*, sehingga diperoleh matriks jarak baru. Perbaikan matriks jarak dilakukan dengan menggunakan metode *complete linkage* dari persamaan (2.1).

Dari hasil perbaikan jarak, diketahui bahwa jarak terdekat terjadi antara *cluster* (d,f) dan (c), sehingga pada tahap selanjutnya kedua *cluster* tersebut digabung. Selanjutnya dilakukan proses yang sama sampai diperoleh satu *cluster* yang berisi semua negara. Proses analisis *cluster* dapat diilustrasikan ke dalam bentuk dendogram. Berikut ini adalah dendogram pengelompokan negara di dunia berdasarkan data runtun waktu realisasi penanaman modal asing di Indonesia tahun 2000-2017:



Gambar 1. Dendogram Analisis *Cluster* Hierarki pada Jarak *Euclidean*.

Selanjutnya akan ditentukan banyak *cluster* optimum dengan perhitungan koefisien *silhouette*. Dengan koefisien tersebut banyak *cluster* optimum ditentukan berdasarkan nilai koefisien *silhouette* yang terbesar. Perhitungan *silhouette* hanya akan dilakukan untuk *cluster* sebanyak  $k = 2$  hingga  $k = 4$  *cluster*. Untuk menda-

patkan nilai *silhouette* suatu  $k$  tertentu, terlebih dahulu dihitung nilai *silhouette* untuk masing-masing negara berdasarkan pengelompokan yang terjadi. Berikut adalah perhitungan koefisien *silhouette* untuk negara USA (objek a) jika banyak *cluster* yang ditentukan  $k = 2$ , yaitu:

$$\begin{aligned} a(a) &= \frac{d_{a,b} + d_{a,c} + \dots + d_{a,q}}{15} \\ &= \frac{4091.5703 + 4173.6744 + \dots + 3458.0713}{15} = 3932.5040 \\ b(a) &= d_{a,o} = 15829,9328, \end{aligned}$$

sehingga diperoleh nilai  $S(a)$  adalah:

$$\begin{aligned} S(a) &= \frac{b(a) - a(a)}{\max\{a(a), b(a)\}} \\ &= \frac{15829.9328 - 3932.5040}{\max\{3932.5040, 15829.9328\}} \\ &= 0.7516 \end{aligned}$$

Dari perhitungan diatas, diperoleh nilai koefisien *silhouette* pada negara USA untuk  $k = 2$  adalah 0.7516. Untuk memperoleh nilai koefisien *silhouette* pada masing-masing *cluster* yang terbentuk, dengan cara menghitung nilai rata-rata koefisien *silhouette* dari seluruh objek yang berada pada *cluster* yang sama. Hasil koefisien *silhouette* keseluruhan negara dapat dilihat pada Tabel 1.

Tabel 1. Koefisien *Silhouette* pada Jarak *Euclidean*

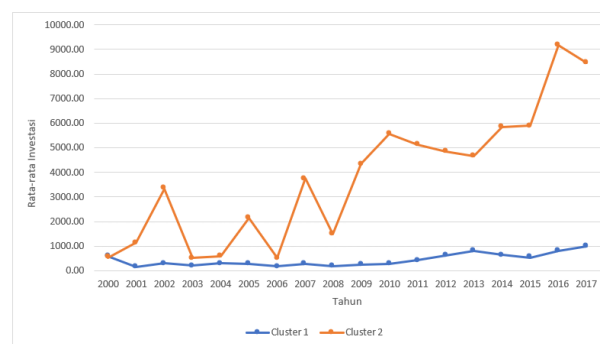
Negara	k=2	k=3	k=4
USA	0.7516	0.4778	0.2712
Kanada	0.8718	0.8189	0.6840
Belgia	0.8717	0.8188	0.6867
Denmark	0.8685	0.8146	0.6732
Perancis	0.8734	0.8202	0.6817
Italia	0.8710	0.8187	0.6833
Belanda	0.7619	0.5269	0.1213
Jerman	0.8608	0.7924	0.5701
Inggris	0.7040	0.4264	0
Swiss	0.8708	0.8108	0.6659
Jepang	0.1551	0	0
Korea Selatan	0.7469	0.4459	0.1284
Hongkong	0.7624	0.5496	0.3403
Taiwan	0.8649	0.8012	0.6239
Singapura	0	0	0
India	0.8735	0.8213	0.6875
Australia	0.8626	0.7918	0.5972
<b>Rata-rata</b>	<b>0.7395</b>	<b>0.6197</b>	<b>0.4362</b>

Berdasarkan Tabel 1 dilihat bahwa semakin banyak jumlah *cluster* maka nilai koefisien *silhouette* semakin rendah. Hal tersebut mengindikasikan bahwa semakin banyak jumlah *cluster* yang terbentuk maka hasil pengelompokan menjadi kurang optimal. Dengan membentuk *cluster* sebanyak 2 diperoleh hasil koefisien *silhouette* dengan rata-rata 0.7395. Berdasarkan kriteria koefisien *silhouette* dapat dikatakan bahwa pengelompokan dengan membentuk 2 *cluster* pada jarak *euclidean* dapat dikategorikan adanya struktur *cluster* yang kuat. Berikut daftar anggota-anggota yang terbentuk sebanyak 2 *cluster*.

Tabel 2. Anggota dari *Cluster* yang Terbentuk dengan Jarak *Euclidean*

<i>Cluster</i>	Anggota
<i>Cluster 1</i>	USA, Kanada, Belgia, Denmark, Perancis, Italia, Belanda, Jerman, Inggris, Swiss, Jepang, Korea Selatan, Hongkong, Taiwan, India, Australia.
<i>Cluster 2</i>	Singapura

Selanjutnya akan dianalisis karakteristik investasi negara di setiap *cluster* yang terbentuk. Karakteristik *cluster* dilihat berdasarkan rata-rata investasi negara asing di Indonesia untuk setiap tahun, dari tahun 2000-2017 sehingga dapat dilihat pengelompokan negara yang besar investasinya memiliki kesamaan selama beberapa tahun tersebut. Pada Gambar 2 ditampilkan pola rata-rata investasi negara.



Gambar 2. Pola Investasi Negara pada *Cluster 1* dan *Cluster 2* Menggunakan Jarak *Euclidean*

Berdasarkan Gambar 2, terlihat bahwa *cluster 1* memiliki pola pergerakan dan besar realisasi investasi yang sangat berbeda dengan *cluster 2*. Untuk *cluster 1* terlihat bahwa besarnya realisasi investasinya sepanjang tahun 2000-2017 berkisar antara 0-1000 juta dolar Amerika dan pergerakan besarnya investasi dari tahun ke tahun cenderung stabil atau tidak mengalami perubahan yang signifikan. Pada tahun 2013 rata-rata investasi meningkat, begitu juga dari tahun 2015 hingga tahun 2017.

Untuk *cluster 2* memiliki bentuk grafik yang berbeda. Pada *cluster 2* terlihat bahwa besarnya realisasi investasinya sepanjang tahun 2000-2017 berfluktuasi serta perubahannya sangat signifikan. Mulai tahun 2008 besarnya investasi tersebut cenderung meningkat sampai saat ini. Pada tahun 2011-2013 terjadi penurunan, namun naik drastis pada tahun selanjutnya, serta mengalami sedikit penurunan pada tahun 2017.

## 5. Kesimpulan

Berdasarkan teknik analisis data yang telah dilakukan diperoleh hasil pengelompokan negara asing sebanyak 2 *cluster* sebagai *cluster* optimum berdasarkan nilai koefisien *silhouette*. Hasil *cluster* pertama terdiri dari 16 negara yaitu Jerman, Swiss, Kanada, India, Belgia, Denmark, Italia, Perancis, Taiwan, Australia, Inggris, Belanda, Hongkong, USA, Jepang dan Korea Selatan. Hasil *cluster* kedua hanya terdiri dari 1 negara yaitu negara Singapura. Pada *cluster 1* besarnya realisasi investasi asing sepanjang tahun 2000-2017 cenderung hampir sama besar. Pada *cluster 2* besarnya realisasi investasi asing sepanjang tahun 2000-2017 lebih besar dibandingkan dengan *cluster 1*.

## 6. Ucapan Terima kasih

Terimakasih kepada Izzati Rahmi HG, M.Si, Dr. Ferra Yanuar, dan Dr. Arrival Rince Putri yang telah memberikan kritik dan saran sehingga tulisan ini menjadi lebih baik.

## Daftar Pustaka

- [1] Aghabozorgi, S., A. S. Shirkorshidi and T. Y. Wah. 2015. Time-series clustering, A decade review. *Information Systems*. **53**: 16 – 38
- [2] Devianto, D. Maiyastri, and D.R. Fadhillah. 2015. Time series modeling for risk of stock price with value at risk computation. *Applied Mathematical Sciences*. **9**(56): 2779 – 2780
- [3] Johnson, R.A. and D.W. Wichern. 1998. *Applied Multivariate Statistical Analysis, Fourth Edition*. Prentice Hall Inc, New Jersey
- [4] Kaufman, L. dan Rousseeuw, P. J. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley, New York
- [5] Martalena, M.Malinda. 2011. *Pengantar Pasar Modal*. Penerbit ANDI, Yogyakarta
- [6] Rencher, A. C. 2002. *Methods of Multivariate Analysis, Second Edition*. Wiley-Interscience, United States of America
- [7] Sartono, B. 2003. *Analisis Peubah Ganda*. IPB Press, Bogor
- [8] Wei, W. W. S. 2006. *Time Series Analysis: Univariate and Multivariate Methods, Second Edition*. Addison Wesley, New York