

**PENERAPAN ANALISIS CLUSTER PADA DATA INDEKS
PEMBANGUNAN TEKNOLOGI INFORMASI DAN
KOMUNIKASI DI INDONESIA**

TUGAS AKHIR

Untuk Memenuhi Sebagian Persyaratan Memperoleh Gelar Ahli Madya



Oleh
MONICA AMIN FEBRIYETI
NIM. 20037041/2020

**PROGRAM STUDI DIPLOMA III STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2023**

PERSETUJUAN TUGAS AKHIR

**PENERAPAN ANALISIS CLUSTER PADA DATA INDEKS
PEMBANGUNAN TEKNOLOGI INFORMASI DAN KOMUNIKASI
DI INDONESIA**

Nama : Monica Amin Febriyeti
NIM : 20037041
Program Studi : D3 Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 02 November 2023

Mengetahui:
Kepala Departemen Statistika



Dr. Yenni Kurniawati, S.Si., M.Si
NIP. 198402232010122005

Disetujui Oleh:
Pembimbing



Dr. Dony Permana, M.Si
NIP. 197501272006041001

PENGESAHAN LULUS UJIAN TUGAS AKHIR




Nama : Monica Amin Febriyeti
NIM : 20037041
Program Studi : D3 Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

PENERAPAN ANALISIS CLUSTER PADA DATA INDEKS PEMBANGUNAN TEKNOLOGI INFORMASI DAN KOMUNIKASI DI INDONESIA

Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Tugas Akhir
Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Padang

Padang, 02 November 2023

Tim Penguji

	Nama	Tanda Tangan
Ketua	: Dr. Dony Permana, M.Si	
Anggota	: Dodi Vionanda, M.Si., Ph.D	
Anggota	: Fadhilah Fitri, S.Si., M.Stat	

SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertandatangan di bawah ini:

Nama : Monica Amin Febriyeti
NIM : 20037041
Program Studi : D3 Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa, skripsi saya dengan judul **"Penerapan Analisis Cluster pada Data Indeks Pembangunan Teknologi Informasi dan Komunikasi di Indonesia"** adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam tradisi keilmuan.

Apabila suatu saat terbukti saya melakukan plagiat maka saya bersedia diproses dan menerima sanksi akademis maupun hukum sesuai dengan hukum dan ketentuan yang berlaku, baik di institusi UNP maupun di masyarakat dan negara.

Demikian pernyataan ini saya buat dengan penuh kesadaran rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Diketahui oleh,
Kepala Departemen Statistika,



Dr. Yenni Kurniawati, S.Si., M.Si
NIP. 19840223201012205

Saya yang menyatakan,



METERAI TEMPEL
JEF AKX795900661

Monica Amin Febriyeti
NIM. 20037041

Penerapan Analisis Cluster pada Data Indeks Pembangunan Teknologi Informasi dan Komunikasi di Indonesia

Monica Amin Febriyeti

ABSTRAK

Seiring dengan perkembangan zaman Teknologi Informasi dan Komunikasi (TIK) sangat pesat perkembangannya di seluruh dunia. Perkembangan TIK sangat perlu diperhatikan serta diperhitungkan untuk dapat melihat bagaimana tingkat pembangunan TIK di suatu wilayah. Indeks Pembangunan Teknologi Informasi dan Komunikasi (IP-TIK) digunakan untuk memperhitungkan tingkat pembangunan TIK. Salah satu penyebab dari permasalahan tingkat pembangunan TIK adalah adanya kesenjangan digital dan belum tepat sasaran kebijakan yang dilakukan oleh pemerintah serta belum meratanya pembangunan TIK di setiap provinsi yang ada di Indonesia. Langkah awal atau solusi yang dapat dilakukan oleh pemerintah adalah mengelompokkan provinsi di Indonesia berdasarkan subindeks dari IP-TIK dengan menggunakan penerapan analisis *cluster*. Analisis *cluster* merupakan teknik pengelompokan objek-objek berdasarkan tingkat kemiripan karakteristik yang dimiliki, dimana tingkat kemiripan tersebut diukur menggunakan ukuran kedekatan jarak antar objek yaitu jarak *Mahalanobis* dan metode pengelompokan *average linkage*.

Jenis penelitian yang digunakan adalah penelitian terapan, yang datanya bersumber dari Publikasi Badan Pusat Statistik Republik Indonesia tahun 2022. Variabel yang digunakan adalah subindeks dari IP-TIK, yaitu *ACCES*, *USE* dan *SKILL* dengan objek yaitu 34 Provinsi di Indonesia tahun 2021.

Hasil analisis *cluster* yang diperoleh yaitu terbentuk menjadi 3 *cluster* dengan memiliki nilai *silhouette* sebesar 0,60 yang menandakan bahwa jumlah *cluster* yang terbentuk termasuk kedalam kriteria struktur yang baik. *Cluster* 1 terdiri dari Provinsi DKI Jakarta dan Provinsi DI Yogyakarta yang termasuk kedalam kategori IP-TIK tinggi. *Cluster* 2 terdiri dari 31 Provinsi yang termasuk kedalam kategori IP-TIK sedang dan *cluster* 3 terdiri dari Provinsi Papua yang termasuk kedalam kategori IP-TIK rendah.

Kata Kunci: Analisis *Cluster*, IP-TIK, Jarak *Mahalanobis*, Metode *Average Linkage*, *Silhouette*, TIK

Application of Cluster Analysis to Information and Communication Technology Development Index Data in Indonesia

Monica Amin Febriyeti

ABSTRACT

Along with the development of the era, Information and Communication Technology (ICT) is developing very rapidly throughout the world. The development of ICT really needs to be paid attention to and calculated to be able to see the level of ICT development in a region. The Information and Communication Technology Development Index (IP-TIK) is used to calculate the level of ICT development. One of the causes of the problem with the level of ICT development is the existence of a digital divide and the lack of targeted policies implemented by the government and the uneven distribution of ICT development in every province in Indonesia. The initial step or solution that can be taken by the government is to group provinces in Indonesia based on sub-indices of IP-TIK using the application of cluster analysis. Cluster analysis is a technique for grouping objects based on the level of similarity of their characteristics, where the level of similarity is measured using a measure of closeness between objects, namely the Mahalanobis distance and the average linkage grouping method.

The type of research used is applied research, the data of which is sourced from the publication of the Central Statistics Agency of the Republic of Indonesia in 2022. The variables used are sub-indices of IP-TIK, namely ACCES, USE and SKILL with objects namely 34 Provinces in Indonesia in 2021.

The results of the cluster analysis obtained were formed into 3 clusters with a silhouette value of 0.60, which indicates that the number of clusters formed falls within the criteria for a good structure. Cluster 1 consists of DKI Jakarta Province and DI Yogyakarta Province which are included in the high IP-ICT category. Cluster 2 consists of 31 provinces which are included in the IP-category Medium ICT and cluster 3 consists of Papua Province which is included in the low IP-ICT category.

Keywords: Cluster Analysis, IP-TIK, Mahalanobis Distance, Average Linkage Method, Silhouette, ICT

KATA PENGANTAR

Puji dan syukur penulis ucapkan atas kehadiran Allah SWT, karena berkat rahmat dan hidayah-Nya penulis dapat menyelesaikan Tugas Akhir yang berjudul **“Penerapan Analisis *Cluster* pada Data Indeks Pembangunan Teknologi Informasi dan Komunikasi di Indonesia”**. Penulisan Tugas Akhir ini merupakan salah satu syarat penyelesaian pendidikan pada Program Studi Diploma III Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang.

Penulisan dan penyusunan Tugas Akhir ini tidak terlepas dari bantuan, bimbingan serta arahan dari berbagai pihak. Pada kesempatan kali ini penulis ingin mengucapkan terima kasih yang sedalam-dalamnya kepada:

1. Bapak Dr. Dony Permana, S.Si, M.Si., pembimbing dan penasehat akademik yang telah banyak memberikan bimbingan, saran, dan arahan demi selesainya Tugas Akhir ini.
2. Bapak Dodi Vionanda, M.Si., Ph.D., Koordinator Program Studi Statistika DIII Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang selaku dosen pembahas Tugas Akhir.
3. Ibu Fadhilah Fitri, S.Si., M.Stat, dosen pembahas Tugas Akhir.
4. Ibu Dr. Yenni Kurniawati, M. Si, Kepala Departemen Statistika sekaligus Koordinator Program Studi Statistika S1 Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang.
5. Bapak/Ibu dosen, staf pengajar serta karyawan Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang.

6. Kedua orang tua, Bapak M. Amin dan Ibu Eti Susanti serta keluarga besar yang tiada hentinya selalu memberikan do'a, dukungan, semangat serta nasihat demi kelancaran Tugas Akhir ini.

7. Semua sahabat yang terlibat dalam penyelesaian Tugas Akhir ini.

Semoga semua bantuan, bimbingan serta arahan tersebut dibalas oleh Allah SWT dan dijadikan sebagai amal ibadah. Penulis menyadari bahwa dalam penulisan serta penyusunan Tugas Akhir ini masih jauh dari kata sempurna. Oleh karena itu, penulis dengan sepenuhnya akan menerima kritik dan saran yang membangun dari pembaca untuk penyempurnaan Tugas Akhir ini. Penulis berharap semoga Tugas Akhir ini dapat bermanfaat bagi penulis dan semua pihak yang membacanya.

Aamiin Ya Rabbal 'Alamiin.

Padang, 2 November 2023

Monica Amin Febriyeti

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR	iii
DAFTAR ISI.....	v
DAFTAR TABEL.....	vii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN.....	ix
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Batasan Masalah.....	8
C. Rumusan Masalah	8
D. Tujuan Penelitian	8
E. Manfaat Penelitian	8
BAB II KAJIAN TEORI.....	10
A. Teknologi Informasi dan Komunikasi (TIK)	10
B. Indeks Pembangunan Teknologi Informasi dan Komunikasi	11
C. Uji Multikolinearitas	18
D. Analisis Cluster	19
BAB III METODOLOGI PENELITIAN.....	31
A. Jenis Penelitian.....	31
B. Jenis dan Sumber Data	31
C. Variabel Penelitian dan Struktur Data.....	31
D. Teknik Analisis Data.....	32
BAB IV HASIL DAN PEMBAHASAN	34
A. Deskripsi Data.....	34
B. Uji Multikolinearitas	40
C. Analisis Data	41
D. Pembahasan.....	47
BAB V KESIMPULAN DAN SARAN.....	49
A. Kesimpulan	49
B. Saran	50

DAFTAR PUSTAKA	51
LAMPIRAN.....	53

DAFTAR TABEL

TABEL	Halaman
1. Nilai dan Peringkat IP-TIK di Beberapa Negara Tahun 2015-2016.....	1
2. Multikolinearitas antar Variabel <i>ACCES</i> (Akses dan Infrastruktur TIK), <i>USE</i> (Penggunaan TIK) dan <i>SKILL</i> (Keahlian TIK)	6
3. Penyusun IP-TIK.....	12
4. Interpretasi Harga r	18
5. Kriteria pengukuran <i>Silhouette Coefficient</i>	30
6. Struktur Data Penelitian	32
7. Uji Multikolineritas Berdasarkan Nilai VIF	41
8. Pembentukan <i>Cluster</i>	44
9. Indeks Validitas Banyaknya <i>Cluster</i>	45
10. Rata-rata Variabel Setiap <i>Cluster</i> Berdasarkan Subindeks dari IP-TIK	45

DAFTAR GAMBAR

GAMBAR	Halaman
1. Perkembangan IP-TIK Indonesia Tahun 2015-2021	3
2. Nilai IP-TIK Provinsi di Indonesia Tahun 2020-2021.....	4
3. Contoh Dendogram	28
5. <i>Boxplot</i> Variabel <i>ACCES</i>	34
6. <i>Boxplot</i> Variabel <i>USE</i>	35
7. <i>Boxplot</i> Variabel <i>SKILL</i>	36
8. <i>Bar Chart</i> Nilai <i>ACCES</i> di Indonesia Tahun 2021	37
9. <i>Bar Chart</i> Nilai <i>USE</i> di Indonesia Tahun 2021	38
10. <i>Bar Chart</i> Nilai <i>SKILL</i> di Indonesia Tahun 2021	39
11. Dendogram Metode <i>Average Linkage</i>	43

DAFTAR LAMPIRAN

LAMPIRAN	Halaman
1. Data Subindeks Indeks Pembangunan Teknologi Informasi dan Komunikasi tahun 2021 di Indonesia.....	53
2. Syntax Python Membuat <i>Boxplot</i>	55
3. Output Python Matriks Variansi-Kovariansi	55
4. Cara Manual Menghitung Matriks Variansi-Kovariansi.....	55
5. Output Python Invers Matriks Variansi-Kovariansi	57
6. Output Python Jarak <i>Mahalanobis</i>	57
7. Cara Manual Menghitung Jarak <i>Mahalanobis</i>	58
8. Output Python Metode <i>Average Linkage</i>	60
9. Output Python Menghitung Indeks Validasi <i>Silhoutte</i>	61
10. Rata-rata <i>Cluster</i>	62
11. Perhitungan Keseluruhan Menggunakan <i>Software Python</i>	64

BAB I PENDAHULUAN

A. Latar Belakang

Pada saat sekarang ini teknologi informasi dan komunikasi sangat pesat perkembangannya di seluruh dunia. Permintaan akan jasa teknologi informasi dan komunikasi sangat diperlukan untuk memenuhi berbagai kebutuhan layanan, seperti dalam hal pendidikan, *e-commerce*, *game online*, *streaming*, komunikasi dan lain sebagainya. Keberadaan akan teknologi informasi dan komunikasi telah mengambil peran penting dalam menunjang pembangunan masyarakat, pemerintah dan bangsa di suatu negara. Seperti halnya di Indonesia, Indeks Pembangunan Teknologi Informasi dan Komunikasi (IP-TIK) sangat penting sebagai ukuran standar tingkat pembangunan TIK di suatu wilayah sehingga dapat dibandingkan antar waktu dan antar wilayah (Indeks Pembangunan TIK, 2022).

Pembangunan TIK di Indonesia mengalami peningkatan dari tahun 2015 ke tahun 2016, dimana pada tahun 2015 IP-TIK Indonesia menduduki peringkat ke-114 dari 175 negara menjadi peringkat ke-111 dari 176 negara (Measuring Information Society ITU, 2016). Jika dibandingkan dengan negara ASEAN lainnya, posisi Indonesia berada di atas Kamboja, Timor Leste, dan Myanmar. Tabel 1 memperlihatkan nilai dan peringkat IP-TIK beberapa negara di dunia.

Tabel 1. Nilai dan Peringkat IP-TIK di Beberapa Negara Tahun 2015-2016

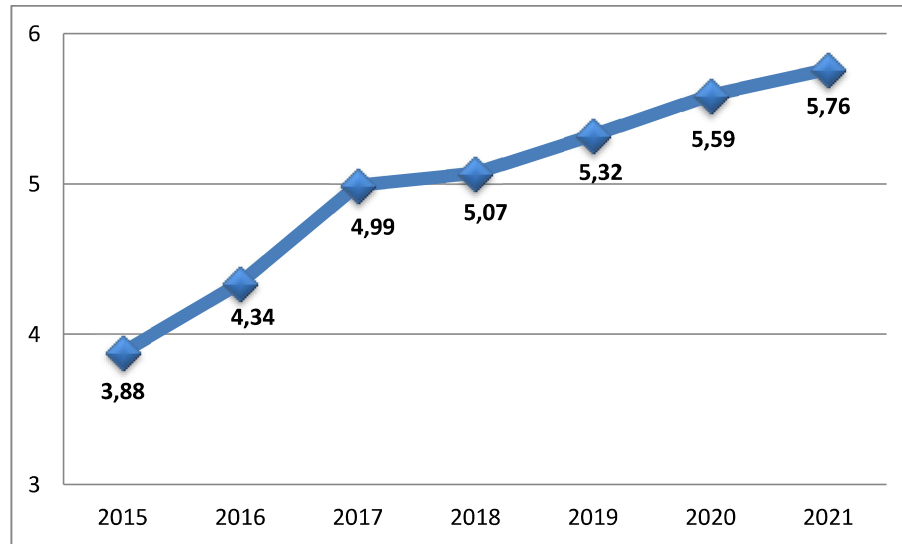
Negara	Indeks Pembangunan Teknologi Informasi dan Komunikasi			
	2015	Peringkat	2016	Peringkat
Korea Selatan	8,80	1	8,85	2
Islandia	8,78	2	8,98	1
Denmark	8,68	3	8,71	4

Swiss	8,66	4	8,74	3
Inggris	8,53	5	8,65	5
Jepang	8,32	11	8,43	10
Australia	8,08	16	8,24	14
Singapura	7,85	20	8,05	18
Malaysia	6,22	62	6,38	63
Brunei Darussalam	6,56	54	6,75	53
Thailand	5,31	79	5,67	78
Vietnam	4,18	108	4,43	108
Filipina	4,52	100	4,67	101
Indonesia	3,85	114	4,33	111
Kamboja	3,24	128	3,28	128
Timor Leste	3,11	127	3,57	122
Myanmar	2,59	140	3,00	135

Indonesia merupakan negara kepulauan yang terdiri dari berbagai provinsi, sehingga dengan adanya IP-TIK dapat dijadikan sebagai alat pengukuran yang penting dalam mengevaluasi tingkat kemajuan teknologi di suatu provinsi baik itu dalam pemanfaatan teknologi, akses informasi, dan kemampuan komunikasi.

Dalam era globalisasi yang didorong oleh perkembangan teknologi, negara-negara dan daerah perlu mengembangkan strategi yang komprehensif untuk meningkatkan IP-TIK mereka. Pembangunan infrastruktur teknologi yang kuat menjadi kunci untuk meningkatkan konektivitas dan aksesibilitas bagi semua lapisan masyarakat. Selain itu, penting juga untuk memberdayakan masyarakat dengan pengetahuan dan keterampilan digital agar dapat memanfaatkan teknologi secara efektif dan berpartisipasi dalam ekonomi digital.

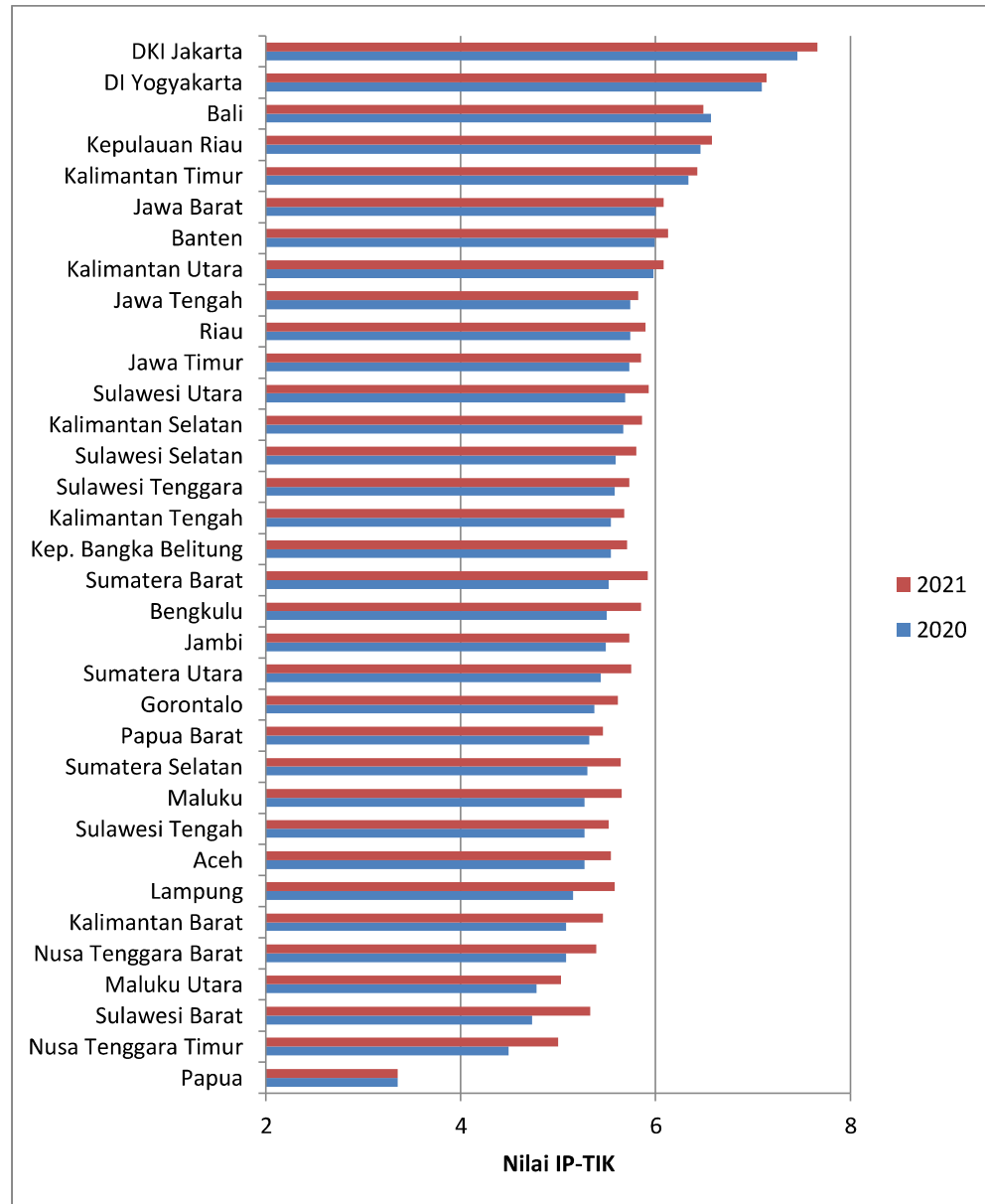
Menurut Badan Pusat Statistik (BPS, 2022) IP-TIK Indonesia selalu mengalami peningkatan dari tahun ke tahun dengan skala 0-10. Gambar 1 memperlihatkan kenaikan IP-TIK Indonesia.



Gambar 1. Perkembangan IP-TIK Indonesia Tahun 2015-2021

Berdasarkan Gambar 1 dapat dilihat bahwa IP-TIK Indonesia terus mengalami peningkatan di setiap tahunnya, dimana yang awalnya pada tahun 2015 sebesar 3,88 meningkat menjadi sebesar 5,76 pada tahun 2021.

Nilai IP-TIK setiap provinsi di Indonesia dari tahun ke tahun juga mengalami peningkatan. Pada tahun 2021 DKI Jakarta menjadi provinsi dengan IP-TIK tertinggi yaitu sebesar 7,66, sedangkan provinsi dengan IP-TIK terendah adalah Papua, yaitu sebesar 3,35. Gambar 2 memperlihatkan peningkatan nilai IP-TIK seluruh provinsi di Indonesia dari tahun 2020-2021.



Gambar 2. Nilai IP-TIK Provinsi di Indonesia Tahun 2020-2021

Berdasarkan Gambar 2 dapat dilihat bahwa hampir seluruh provinsi di Indonesia dari tahun 2020-2021 mengalami peningkatan nilai IP-TIK. Terdapat 2 provinsi yang tidak mengalami peningkatan nilai IP-TIK, yaitu Provinsi Bali pada tahun 2020 memiliki nilai IP-TIK sebesar 6,57 mengalami penurunan pada tahun 2021 menjadi 6,49 dan Provinsi Papua pada tahun 2020-2021 memiliki nilai IP-

TIK yang sama, yaitu sebesar 3,35. Provinsi DKI Jakarta merupakan provinsi yang memiliki nilai IP-TIK tertinggi yaitu berada direntang nilai 7,51-10,00, hal ini menandakan bahwa IP-TIK di Provinsi DKI Jakarta berada di kategori tinggi. Kategori IP-TIK rendah berada di Provinsi Papua, dimana nilai IP-TIK hanya berada direntang nilai 2,51-5,00.

Pada tahun 2020 Suharni melakukan pengelompokan wilayah berdasarkan IP-TIK ke dalam 3 *cluster* menggunakan metode *K-Means*, sehingga diperoleh *cluster* rendah (C1) terdiri atas 4 Provinsi yaitu Nusa Tenggara Timur, Sulawesi Barat, Maluku Utara, dan Papua, *cluster* sedang (C2) terdiri atas 22 Provinsi dan *cluster* tinggi (C3) terdiri dari 8 Provinsi yaitu Kepulauan Riau, DKI Jakarta, Jawa Barat, DI Yogyakarta, Banten, Bali, Kalimantan Timur dan Kalimantan Utara.

Adanya perbedaan nilai IP-TIK antar provinsi dikarenakan adanya kesenjangan digital, dimana subindeks-subindeks yang mendasari penghitungan IP-TIK tidak terpenuhi antar provinsi. Pemantauan berkelanjutan terhadap perkembangan TIK sangat penting bagi para pengambil kebijakan. Mengingat dampak potensial dari penggunaan TIK pada pembangunan sosial dan ekonomi, Namun suatu kebijakan harus berdasarkan pada bukti dan fakta yang terukur serta indikator yang dapat diperbandingkan. Indikator ini digunakan untuk membandingkan pencapaian TIK masing-masing provinsi dan menjadi tolak ukur yang penting untuk menilai daya saing baik di tingkat regional maupun global, sehingga akan berdampak pada peningkatan TIK di tingkat nasional. Indikator tersebut terbagi ke dalam 3 subindeks, dengan adanya subindeks ini dapat dilakukan suatu pengelompokan provinsi menggunakan analisis *cluster*.

Analisis *cluster* termasuk kedalam teknik analisis multivariat yang tujuan utamanya adalah untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristik yang dimilikinya. Kemiripan karakteristik tersebut dapat ditentukan dari kedekatan jarak, ukuran jarak *Euclid* atau ukuran jarak *Mahalanobis*. Jarak *Euclid* digunakan ketika tidak ada terjadi korelasi atau multikolinearitas antar objek, sedangkan jarak *Mahalanobis* digunakan ketika terdapat data yang berkorelasi atau multikolinearitas (Rahmawati, 2007). Tabel 2 memperlihatkan hubungan antar variabel pada penelitian ini yang diperoleh menggunakan bantuan *software* Minitab 16.

Tabel 2. Multikolinearitas antar Variabel *ACCES* (Akses dan Infrastruktur TIK), *USE* (Penggunaan TIK) dan *SKILL* (Keahlian TIK)

	<i>ACCES</i>	<i>USE</i>
<i>USE</i>	0,876 0,000	
<i>SKILL</i>	0,539 0,001	0,357 0,038

Berdasarkan Tabel 2, dapat dilihat bahwa ketiga variabel tersebut memiliki multikolinearitas antar satu variabel dengan variabel yang lainnya. Hal ini ditandai dengan nilai korelasi yang diperoleh antar variabel, dimana untuk variabel *ACCES* dengan *USE* berkorelasi sangat tinggi yang nilainya sebesar 0,876. Variabel *ACCES* dengan *SKILL* cukup berkorelasi, dimana hanya memiliki nilai sebesar 0,539. Sedangkan variabel *USE* dengan *SKILL* berkorelasi rendah.

Terdapat tiga metode yang umum digunakan untuk pengelompokkan objek dalam analisis *cluster*, yaitu metode grafik, metode hirarki dan metode non hirarki. Metode grafik merupakan metode yang hanya efektif untuk data yang tidak terlalu banyak objeknya, terdiri atas plot profil, plot Andrew dan plot

Andrew termodifikasi. Metode hirarki merupakan metode yang digunakan untuk mengelompokkan objek secara terstruktur berdasarkan kemiripan karakteristiknya dan kelompok yang diinginkan belum diketahui jumlahnya, dimana metode ini terdiri atas metode yang paling sering digunakan dalam pengelompokkan yaitu *single linkage*, *complete linkage*, dan *average linkage*. Sedangkan metode non hirarki merupakan metode yang digunakan untuk mengelompokkan pengamatan dimana jumlah kelompok yang akan dibentuk dapat ditentukan terlebih dahulu dan biasanya digunakan untuk jumlah objek yang lebih besar, salah satu metodenya yaitu *K-Means* (Mattjik, 2011: 198-214).

Proses pengelompokkan pada penelitian ini menggunakan metode *average linkage* atau pautan rata-rata, dengan menggunakan jarak *Mahalanobis* sebagai ukuran jarak antar objek (ukuran kemiripan karakteristik). Menurut Aprilia (2016) metode ini lebih stabil dibandingkan dengan metode lainnya yaitu metode *single linkage* dan *complete linkage*. Berdasarkan penelitian yang dilakukan oleh Wirdayani (2020) juga menunjukkan bahwa metode *average linkage* merupakan analisis *cluster* terbaik dalam pengelompokkan kabupaten/kota di Indonesia terhadap indikator kemiskinan multidimensi. Penelitian yang sama dilakukan di Provinsi Daerah Istimewa Yogyakarta pada tahun 2016 untuk mengelompokkan penderita disabilitas mental, juga mengatakan bahwa metode *average linkage* merupakan metode *cluster* yang terbaik (Widodo *et al.*, 2018).

Berdasarkan permasalahan yang telah dijabarkan, penelitian ini diberi judul **“Penerapan Analisis Cluster pada Data Indeks Pembangunan Teknologi Informasi dan Komunikasi di Indonesia”**

B. Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah:

1. Mengelompokkan provinsi yang ada di Indonesia pada tahun 2021
2. Menggunakan metode *average linkage* berdasarkan subindeks dari Indeks Pembangunan Teknologi Informasi dan Komunikasi (IP-TIK), yang terdiri dari subindeks *ACCES*, subindeks *USE*, dan subindeks *SKILL*.

C. Rumusan Masalah

Adapun rumusan masalah yang akan dibahas dalam penelitian ini adalah:

1. Bagaimana hasil pengelompokkan provinsi di Indonesia berdasarkan subindeks dari IP-TIK menggunakan analisis *cluster*?
2. Bagaimana ciri kelompok provinsi di Indonesia berdasarkan subindeks dari IP-TIK yang terbentuk menggunakan analisis *cluster*?

D. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Untuk mengetahui pengelompokkan provinsi di Indonesia berdasarkan subindeks dari IP-TIK menggunakan analisis *cluster*
2. Untuk mengetahui ciri kelompok provinsi di Indonesia berdasarkan subindeks dari IP-TIK yang terbentuk menggunakan analisis *cluster*

E. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat seperti sebagai berikut:

1. Bagi penulis, dapat menambah wawasan dan memperdalam pengetahuan dalam mengaplikasikan analisis *cluster* terkhususnya menggunakan metode *average linkage*.

2. Bagi pemerintah, sebagai bahan pertimbangan yang dapat dijadikan pemerintah dalam mengambil sebuah kebijakan mengenai perkembangan IP-TIK diberbagai provinsi yang ada di Indonesia, serta membantu dalam menentukan provinsi mana saja yang lebih diprioritaskan untuk mendapatkan pembangunan dalam hal TIK agar dapat memperkecil kesenjangan digital.
3. Bagi peneliti lainnya, sebagai bahan referensi dalam melakukan penelitian selanjutnya.