



TUGAS AKHIR - SS141501

**PENGELOMPOKAN KABUPATEN/KOTA DI
PROVINSI JAWA TIMUR BERDASARKAN
INDIKATOR KESEHATAN MASYARAKAT
MENGUNAKAN METODE KOHONEN SOM
DAN K-MEANS**

**MARINA MARSUDI PUTRI
NRP 1311 100 028**

**Dosen Pembimbing
Dr. Dra. Kartika Fithriasari, M.Si.**

**JURUSAN STATISTIKA
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2015**



FINAL PROJECT- SS141501

**GROUPING DISTRICTS/CITIES IN EAST JAVA BASED
ON PUBLIC HEALTH INDICATORS USING KOHONEN
SOM AND K-MEANS**

MARINA MARSUDI PUTRI
NRP 1311 100 028

Supervisor
Dr. Dra. Kartika Fithriasari, M.Si.

DEPARTMENT OF STATISTICS
Faculty of Mathematics and Natural Sciences
Sepuluh Nopember Institute of Technology
Surabaya 2015

LEMBAR PENGESAHAN

**PENGELOMPOKAN KABUPATEN/KOTA DI JAWA
TIMUR BERDASARKAN INDIKATOR KESEHATAN
MASYARAKAT MENGGUNAKAN METODE
KOHONEN SOM DAN K-MEANS**

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada
Program Studi S-1 Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :

MARINA MARSUDI PUTRI
NRP. 1311 100 028

Disetujui oleh Pembimbing Tugas Akhir :

Dr. Dra. Kartika Fithriasari, M.Si. ()
NIP. 19691212 199303 2 002

Mengetahui
Ketua Jurusan Statistika FMIPA-ITS




Dr. Muhammad Mashuri, MT
NIP. 19620408 198701 1 001

SURABAYA, JANUARI 2015

**PENGELOMPOKAN KABUPATEN/KOTA DI JAWA TIMUR
BERDASARKAN INDIKATOR KESEHATAN
MASYARAKAT MENGGUNAKAN METODE KOHONEN
SOM DAN K-MEANS**

Nama : Marina Marsudi Putri
NRP : 131110028
Jurusan : Statistika FMIPA – ITS
Dosen Pembimbing : Dr. Dra. Kartika Fithriasari, M.Si.

Abstrak

Kondisi kesehatan masyarakat di Jawa Timur yang tidak homogen akan menyulitkan bagi Tim Pembina Kota Sehat pada saat melakukan pembinaan dan monitoring. Berdasarkan hal tersebut diperlukan pengelompokan kabupaten dan kota di Jawa Timur yang didasarkan pada kemiripan karakteristik kondisi kesehatan masyarakat, sehingga proses pembinaan dan monitoring akan dilakukan berdasarkan kelompok kabupaten dan kota yang terbentuk. Pengelompokan kabupaten/kota di Jawa Timur akan dilakukan menggunakan metode Kohonen SOM kemudian akan dibandingkan dengan hasil pengelompokan metode klasik K-Means, yang merupakan metode pengelompokan yang populer dan sering digunakan, berdasarkan kriteria nilai icdrate (internal cluster dispersion rate). Pada pengujian Bartlett diperoleh kesimpulan bahwa terdapat hubungan atau korelasi di antara variabel penelitian. Setelah dilakukan analisis faktor untuk mereduksi variabel diperoleh 3 faktor baru yang terbentuk. Berdasarkan nilai Pseudo Fstatistics yaitu sebesar 13,819, hasil pengelompokan terbaik adalah menggunakan metode Kohonen SOM dengan jenis topologi hextop. Sedangkan pada metode K-Means nilai Pseudo Fstatistics terbesar yaitu 9,781 ketika digunakan kelompok sebanyak 2 klaster. Perbandingan hasil klaster terbaik berdasarkan nilai icdrate diperoleh kesimpulan bahwa nilai icdrate metode Kohonen SOM yaitu sebesar 0,962 lebih kecil dibandingkan dengan nilai icdrate hasil pengelompokan metode K-Means yaitu sebesar 0,988. Pada hasil pengujian One-way MANOVA diperoleh kesimpulan bahwa pada masing-masing kelompok yang terbentuk memiliki perbedaan.

Kata Kunci—*Icdrate, Indikator Kesehatan Masyarakat, K-means, Kohonen SOM, One-way MANOVA, Pseudo F-statistics*

(Halaman ini sengaja dikosongkan)

GROUPING DISTRICTS/CITIES IN EAST JAVA BASED ON PUBLIC HEALTH INDICATORS USING KOHONEN SOM AND K-MEANS

Name : Marina Marsudi Putri
NRP : 1311100028
Departement : Statistics FMIPA – ITS
Supervisor : Dr. Dra. Kartika Fithriasari, M.Si.

Abstract

Public health conditions in East Java is not homogeneous. That will be difficult for Team evaluation to monitoring about public health condition in that area. Based on the required, grouping of districts and cities in East Java, which is based on the similarity of the characteristics of public health conditions, so that the process of monitoring will be done by district and city groups are formed. The grouping of districts/cities in East Java will done by using Kohonen SOM and then will compared with the results of grouping classic K-Means method, based on the criteria of value icdrate (internal cluster dispersion rate). In the Bartlett test is concluded that there is a relationship or correlation between the study variables. After analyzing the factors to reduce the variables obtained three new factors were formed. Based on the value of the Pseudo Fstatistics that is 13,819, the best result for clustering using Kohonen SOM with hextop topolgy. Beside that the value of Pseudo Fstatistics using K-Means with 2 cluster is 9,871. The result of the comparison from clustering based on icdrate value is Kohonen SOM has value 0,962 is smaller than icdrate value from K-Means that is 0,988. From One-way NAMOVA testing the conclusion is cluster that performed has different characteristics.

Keyword-*Icdrate, K-means, Kohonen SOM, One-way MANOVA, Pseudo F-statistics, Public Health Indicators*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Segala puji bagi Allah SWT, Tuhan semesta alam, sehingga penulis dapat menyelesaikan tugas akhir yang berjudul:

**PENGELOMPOKAN KABUPATEN/KOTA DI
JAWA TIMUR BERDASARKAN INDIKATOR
KESEHATAN MASYARAKAT MENGGUNAKAN
METODE KOHONEN SOM DAN K-MEANS**

Terselesaikannya Tugas Akhir ini tidak terlepas dari dukungan berbagai pihak yang telah memberikan bimbingan dan bantuan pada penulis. Untuk itu pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Ibu Dr. Dra. Kartika Fithriasari M.Si, selaku dosen pembimbing yang dengan sabar membimbing dari awal hingga akhir penyusunan tugas akhir ini dan selalu memberi masukan kepada penulis. Terima kasih banyak, Ibu.
2. Ibu Dra. Wiwiek Setya Winahju, MS selaku dosen wali dan dosen penguji serta Bapak Dr. Brodjol Sutijo Suprih Ulama, M.Si selaku dosen penguji yang telah memberi saran sehingga menjadikan tugas akhir ini lebih baik.
3. Dr. Muhammad Mashuri, MT. selaku Ketua Jurusan Statistika ITS yang telah memberikan fasilitas untuk kelancaran penyelesaian Tugas Akhir ini.
4. Ibu Dra. Lucia Aridinanti, MT. selaku Kaprodi S1 Jurusan Statistika ITS yang juga telah memberikan fasilitas untuk kelancaran penyelesaian Tugas Akhir ini.
5. Ibu, Ayah atas segala do'a, pengorbanan, motivasi, dan kepercayaan yang telah diberikan. Serta adik Bintoro yang selalu mengingatkan tentang *deadline*.
6. Teman seperjuangan Tugas Akhir: Lucky dan Icha, terima kasih sudah menyempatkan dan meluangkan waktu untuk selalu membantu dan memberikan motivasi agar segera menyelesaikan Tugas Akhir ini.
7. Sahabat tercinta Jainap, Nurina, Onya yang sudah bersedia mendengarkan keluh kesah.
8. Dyah, Kyky, Rika, Ipat, Sinta yang selalu memotivasi untuk segera menyelesaikan Tugas Akhir.

9. Seluruh keluarga besar Jurusan Statistika FMIPA Institut Teknologi Sepuluh Nopember Surabaya, khususnya $\Sigma 22$ atas kebersamaan dan kehangatannya.
10. Bu Ui Dinkes yang telah banyak membantu Penulis selama proses pengambilan data. Terima kasih banyak Ibu.
11. Icha dan mbak Meilan, mbak Rizky yang sudah mengajarkan cara-cara menghitung.
12. Serta pihak-pihak lain yang tidak dapat penulis sebutkan satu per satu.

Penulis mengharapkan Tugas Akhir ini dapat memberikan manfaat bagi para pembaca. Penulis menyadari bahwa Tugas Akhir ini masih jauh dari kesempurnaan. Oleh karena itu penulis menerima apabila ada saran dan kritik yang sifatnya membangun guna perbaikan untuk penelitian-penelitian selanjutnya.

Surabaya, Januari 2015

Penulis

DAFTAR ISI

halaman

HALAMAN JUDUL	i
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	
2.1 Uji Bartlett	7
2.2 Analisis Faktor	8
2.3 Jaringan Syarat Tiruan	9
2.4 Kohonen SOM	12
2.5 Metode <i>K-Means</i>	16
2.6 Pemilihan Jumlah Klaster Optimum	18
2.7 Kriteria Perbandingan Klaster Terbaik	19
2.8 Pengujian <i>One-way</i> MANOVA	20
2.9 Indeks Kesehatan Masyarakat	21
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data	25
3.2 Variabel Penelitian	25
3.3 Langkah Analisis	26
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Statistika Deskriptif	31
4.1.1 Statistika Deskriptif Variabel Ekonomi	32

4.1.2	Statistika Deskriptif Variabel Pendidikan	34
4.1.3	Statistika Deskriptif Variabel Epidemiologi	36
4.2	Mereduksi Variabel Penelitian	39
4.3	Pemilihan Klaster Terbaik	42
4.4	Pengelompokan Menggunakan Metode Terbaik	45
BAB V KESIMPULAN DAN SARAN		
5.1	Kesimpulan	53
5.2	Saran	54
DAFTAR PUSTAKA		57
LAMPIRAN		59
BIODATA PENULIS		75

DAFTAR TABEL

	halaman
Tabel 2.1 Istilah dalam Jaringan Syaraf Tiruan	11
Tabel 3.1 Variabel Penelitian	25
Tabel 4.1 Nilai <i>Eigen Value</i> Analisis Faktor.....	40
Tabel 4.2 Nilai <i>Loading Factor</i> untuk Reduksi Variabel	41
Tabel 4.3 Rotasi Nilai <i>Loading Factor</i> untuk Reduksi Variabel	41
Tabel 4.4 Nilai <i>Pseduo F-statistics</i> untuk Metode Kohonen SOM	42
Tabel 4.5 Nilai <i>Icdrate</i> , SSW dan SSB untuk Metode Kohonen SOM.....	44
Tabel 4.6 Nilai Rata-rata Hasil Pengelompokan Metode Terbaik	47
Tabel 4.7 Hasil Pengujian <i>One-way</i> MANOVA pada Klaster ..	50

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

	halaman
Gambar 2.1	Arsitektur Jaringan Kohonen SOM 13
Gambar 2.2	Topologi <i>Gridtop</i> 14
Gambar 2.3	Topologi <i>Hextop</i> 14
Gambar 2.4	Topologi <i>Randtop</i> 14
Gambar 3.1	Diagram Alir Penelitian 29
Gambar 4.1	Grafik PDRB per Kapita Menurut Harga Konstan 32
Gambar 4.2	Grafik Angka Harapan Hidup di Jawa Timur 34
Gambar 4.3	Grafik Angka Melek Huruf di Jawa Timur 35
Gambar 4.4	Grafik Rata-rata Lama Sekolah di Jawa Timur 36
Gambar 4.5	Grafik Prevalensi TBC di Jawa Timur 37
Gambar 4.6	Grafik Prevalensi Malaria di Jawa Timur 38
Gambar 4.7	Arsitektur Jaringan Kohonen SOM Menggunakan 3 Neuron 47

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada tahun 2013 tercatat sebanyak 31 kabupaten/kota di Provinsi Jawa Timur telah mengadopsi program Kota Sehat Nasional, sedangkan 7 kabupaten/kota yang lain akan mengadopsi program Kota Sehat Nasional pada tahun 2014 (Artikel dan Berita Dinas Kesehatan Jawa Timur, 2013). Pemerintah Provinsi Jawa Timur menargetkan pada tahun 2014 seluruh kabupaten atau kota di Jawa Timur telah mengadopsi program Kota Sehat Nasional. Melalui Dinas Kesehatan, Pemerintah Provinsi Jawa Timur akan memfasilitasi kabupaten/kota yang akan dan telah mengadopsi Program Kota Sehat dengan membentuk Tim Pembina Kota Sehat. Tim Pembina Kota Sehat bertugas untuk melakukan monitoring dan evaluasi terhadap kinerja masing-masing kabupaten dan kota dalam bidang kesehatan.

Provinsi Jawa Timur terdiri dari 38 kabupaten/kota yang memiliki karakteristik kondisi kesehatan masyarakat yang berbeda di setiap wilayah. Kondisi kesehatan masyarakat di Jawa Timur yang tidak homogen akan menyulitkan bagi Tim Pembina Kota Sehat pada saat melakukan pembinaan dan monitoring. Oleh karena itu proses pembinaan dan monitoring yang dilakukan perlu disesuaikan dengan karakteristik kondisi kesehatan masyarakat di tiap kabupaten/kota. Namun apabila proses pembinaan dilakukan secara berbeda pada setiap kabupaten/kota akan memerlukan waktu yang lama dan biaya yang tidak sedikit. Berdasarkan hal tersebut diperlukan pengelompokan kabupaten/kota di Jawa Timur yang didasarkan pada kemiripan karakteristik kondisi kesehatan masyarakat, sehingga proses pembinaan dan monitoring akan dilakukan berdasarkan kelompok-kelompok kabupaten/kota yang terbentuk. Indikator kesehatan dapat digunakan untuk mengetahui kondisi kesehatan masyarakat di suatu wilayah, maka dari itu pengelompokan kabupaten/kota di Jawa Timur akan dila-

kukan berdasarkan pada faktor pembentuk indikator kesehatan di tiap wilayah.

Penelitian sebelumnya tentang pengelompokan dalam bidang kesehatan telah dilakukan oleh Shoalina (2014) dengan judul Analisis *Cluster* pada Indikator yang Membentuk Indeks Pembangunan Kesehatan Masyarakat di Kabupaten/Kota Provinsi Jawa Timur menggunakan metode *Ward's*. Nugroho (2014) mengelompokan kabupaten/kota di Jawa Timur berdasarkan indikator kesehatan yang dibagi ke dalam indikator *input*, proses dan *output* menggunakan Analisis Biplot. Ramadhani (2014) melakukan penelitian tentang pengelompokan kecamatan di kota Surabaya berdasarkan derajat kesehatan menggunakan analisis klaster metode *K-Means*. Hasil penelitian yang telah dilakukan oleh Ramadhani (2014) menyatakan bahwa pada kota Surabaya terbentuk tiga kelompok (klaster). Klaster satu yaitu kelompok dengan derajat kesehatan tinggi memiliki karakteristik variabel kualitas pendidikan dan Angka harapan Hidup yang tinggi, tetapi memiliki angka kepemilikan jamban yang rendah. Klaster dua yaitu kelompok dengan derajat kesehatan sedang memiliki karakteristik jumlah penduduk, jumlah kelahiran, jumlah kematian akibat penyakit DBD dan TBC yang cukup tinggi. Klaster tiga yaitu kelompok dengan derajat kesehatan rendah memiliki karakteristik angka partisipasi pendidikan yang rendah namun memiliki jumlah kegiatan penyuluhan kesehatan yang tinggi. Berdasarkan penelitian tersebut dapat disimpulkan bahwa analisis klaster dapat digunakan untuk mengelompokan obyek dan mengetahui karakteristik dari kelompok yang terbentuk.

Selain menggunakan beberapa metode yang telah disebutkan sebelumnya, pengelompokan juga dapat dilakukan menggunakan metode Kohonen SOM (*Self Organizing Map*). Kohonen SOM memiliki keunggulan yaitu dapat mengelompokan data yang mengandung *overlapping* dan mengatasi sifat non linier pada klaster yang terbentuk (Ettaouil, et al., 2012). Keunggulan lain dari metode Kohonen SOM disampaikan oleh Toor dan Singh (2013) dalam sebuah penelitian untuk membandingkan tiga jenis

metode pengelompokan yaitu Kohonen SOM, *K-Means* dan *Hierarchical Clustering Algorithm* (HAC) untuk mengelompokan institusi pendidikan berdasarkan kegiatan pembelajaran dengan kriteria jumlah klaster yang terbentuk, nilai *error rate*, lama waktu pemrosesan dan peta hasil kelompok menggunakan jumlah data yang cukup besar. Penelitian tersebut menghasilkan kesimpulan bahwa ketiga metode masing-masing mengelompokan obyek menjadi 6 klaster dan metode Kohonen SOM memiliki nilai *error rate* terkecil serta waktu pemrosesan tercepat ketika digunakan pada jumlah data dan parameter yang sama. Toor dan Singh juga menyebutkan bahwa metode pengelompokan menggunakan Kohonen SOM akan menghasilkan kesimpulan yang tidak jauh berbeda ketika menggunakan metode *K-Means*.

Indikator kesehatan terdiri atas angka-angka mortalitas, angka-angka morbiditas, indikator status gizi, indikator keadaan lingkungan, indikator perilaku hidup masyarakat, indikator akses dan mutu pelayanan kesehatan, indikator pelayanan kesehatan, indikator sumber daya kesehatan, indikator manajemen kesehatan serta indikator kontribusi sektor-sektor terkait (Dinkes RI, 2010). Gerring, et al (2013) melakukan penelitian untuk mengukur kinerja sistem kesehatan menggunakan indikator kesehatan dengan mempertimbangkan faktor wilayah. Hasil penelitian tersebut menyatakan bahwa tidak semua variabel yang membentuk indikator kesehatan dapat digunakan untuk mengukur tingkat kesehatan di suatu wilayah sebab masing-masing wilayah memiliki karakteristik yang berbeda. Gerring, et al (2013) menyebutkan bahwa variabel yang mempengaruhi indikator kesehatan antara lain yaitu variabel ekonomi, pendidikan dan epidemiologi.

Pada penelitian ini pengelompokan kabupaten/kota di Jawa Timur akan didasarkan pada variabel yang membentuk indikator kesehatan pada hasil penelitian yang telah dilakukan oleh Gerring, et al. pada tahun 2013 menggunakan metode Kohonen SOM. Hal ini dilakukan sebab metode Kohonen SOM dapat mengelompokan data yang mengandung *overlapping* dan mengatasi sifat non linier pada klaster yang terbentuk. Hasil pengelompokan kabu-

paten dan kota menggunakan metode Kohonen SOM akan dibandingkan dengan hasil pengelompokan metode klasik *K-Means*, yang merupakan metode pengelompokan yang populer dan sering digunakan, berdasarkan kriteria nilai *icdrate* (*internal cluster dispersion rate*) yang merupakan tingkat dispersi di dalam klaster.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, masalah yang akan dibahas dalam penelitian ini adalah kondisi kesehatan masyarakat di Jawa Timur yang tidak homogen akan menyulitkan Pemerintah Provinsi Jawa Timur pada saat melakukan pembinaan dan monitoring di setiap kabupaten/kota dalam rangka mengadopsi Program Kota Sehat Nasional. Oleh karena itu akan dilakukan pengelompokan kabupaten/kota di Jawa Timur berdasarkan kemiripan kondisi kesehatan masyarakat sehingga akan memberikan kemudahan bagi Tim Pembina Kota Sehat pada saat melakukan evaluasi dan monitoring. Salah satu metode yang dapat digunakan untuk melakukan pengelompokan adalah metode Kohonen SOM. Hasil pengelompokan kabupaten/kota di Jawa Timur menggunakan metode Kohonen SOM kemudian akan dibandingkan dengan hasil pengelompokan menggunakan metode *K-Means* berdasarkan kriteria nilai *icdrate*.

1.3 Tujuan

Tujuan yang ingin dicapai pada penelitian ini adalah:

1. Mendapatkan jumlah kelompok optimum untuk mengelompokkan kabupaten/kota di Jawa Timur Indikator Kesehatan Masyarakat pada masing-masing metode Kohonen SOM dan *K-Means*.
2. Mendapatkan arsitektur jaringan Kohonen SOM yang sesuai untuk membentuk kelompok kabupaten/kota berdasarkan Indikator Kesehatan Masyarakat.
3. Mengelompokkan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat dengan menggunakan metode Kohonen SOM.

4. Mengelompokkan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat dengan menggunakan metode *K-Means*.
5. Membandingkan hasil pengelompokan menggunakan metode Kohonen SOM dan *K-Means* berdasarkan kriteria kriteria nilai *icdrate*.

1.4 Manfaat

Dengan melakukan penelitian ini diharapkan dapat memperoleh manfaat yaitu memberikan tambahan informasi kepada Pemerintah Daerah Provinsi Jawa Timur tentang kecenderungan indikator kesehatan masyarakat di kelompok kabupaten/kota yang terbentuk. Selain itu hasil penelitian diharapkan dapat juga digunakan sebagai informasi untuk memonitoring kondisi kesehatan di kabupaten/kota dalam rangka mencapai predikat Kota Sehat Nasional.

1.5 Batasan Penelitian

Pada penelitian ini data yang digunakan merupakan variabel pembentuk indikator kesehatan masyarakat di 38 kabupaten/kota yang diperoleh dari hasil Riskesdas dan Badan Pusat Statistik Provinsi Jawa Timur pada tahun 2012.

(Halaman ini sengaja dikosongkan)

BAB II TINJAUAN PUSTAKA

Pada bagian ini akan dibahas mengenai landasan teori yang digunakan untuk mencapai tujuan penelitian. Landasan teori tersebut meliputi pengujian kebebasan variabel penelitian menggunakan Uji Bartlett, analisis faktor, pengertian Jaringan Syaraf Tiruan, metode Kohonen SOM, metode *K-Means*, cara untuk mendapatkan jumlah kluster optimum, kriteria perbandingan kluster terbaik berdasarkan nilai *icdrate*, serta pengujian *One-way* MANOVA.

2.1 Uji Bartlett

Uji Bartlett digunakan untuk mengetahui ada atau tidaknya hubungan antar variabel dalam kasus multivariat. Jika variabel x_1, x_2, \dots, x_p bersifat saling bebas (independen), maka matriks korelasi antar variabel sama dengan matriks identitas. Berdasarkan hal tersebut untuk menguji kebebasan antar variabel hipotesis pengujian yang digunakan pada Uji Bartlett adalah sebagai berikut:

$H_0 : \rho = I$ artinya tidak terdapat hubungan antar variabel

$H_1 : \rho \neq I$ artinya terdapat hubungan antar variabel

Nilai statistik Uji Bartlett dapat ditulis dengan persamaan sebagai berikut:

$$\chi^2 = -(n-1 - \frac{2p+5}{6}) \ln|R| \quad (2.1)$$

Keterangan:

$\ln|R|$: nilai determinan dari matriks korelasi

n : banyaknya observasi

p : banyaknya variabel

Tolak H_0 jika $\chi^2 > \chi^2_{(1/2p)(p-1), \alpha}$ atau jika $p\text{-value} < \alpha$, yang berarti bahwa terdapat hubungan antar variabel.

(Morrison, 1990)

2.2 Analisis Faktor

Analisis faktor merupakan salah satu metode multivariat yang digunakan untuk menganalisis variabel-variabel yang diduga memiliki keterkaitan satu dengan yang lain sehingga keterkaitan tersebut dapat dijelaskan dan dipetakan atau dikelompokkan pada faktor yang tepat. Pada analisis faktor setiap variabel dinyatakan dalam kombinasi linier dari faktor yang mendasari (*underlying factor*). Jumlah varian yang diberikan oleh suatu variabel dengan variabel lainnya disebut dengan *communality*. Kovarian antara variabel yang diuraikan dinyatakan dalam suatu *common factors* ditambah dengan faktor untuk setiap variabel.

Dimisalkan terdapat variabel *random* X dengan komponen sebanyak p , yang memiliki nilai rata-rata μ dan matriks kovarian Σ , maka model faktor X adalah linier tergantung pada beberapa variabel *random* yang tidak teramati, yaitu F_1, F_2, \dots, F_m yang disebut sebagai *common factors* dan ditambahkan sebanyak p variasi $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ yang disebut sebagai *spesifik factors*. Persamaan model analisis faktor dapat ditulis sebagai berikut:

$$\begin{aligned} X_1 - \mu_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1 \\ X_2 - \mu_2 &= \ell_{21}F_1 + \ell_{22}F_2 + \dots + \ell_{2m}F_m + \varepsilon_2 \\ &\vdots \\ X_p - \mu_p &= \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pm}F_m + \varepsilon_p \end{aligned} \quad (2.2)$$

Persamaan (2.2) dapat juga ditulis dalam bentuk matriks sebagai berikut:

$$X - \mu = \underset{(px1)}{L} \underset{(mx1)}{F} + \underset{(px1)}{\varepsilon} \quad (2.3)$$

Keterangan:

F_k = common faktor ke- k

ℓ_{hk} = loading faktor ke- k dan variabel ke- h

μ_h = rata-rata variabel ke- h

ε_h = spesifik faktor ke- h

$i = 1, 2, \dots, n$ adalah banyak observasi

$k = 1, 2, \dots, m$ adalah banyaknya common faktor

$h = 1, 2, \dots, p$ adalah banyaknya variabel

Asumsi yang harus dipenuhi oleh vektor *common factor* (F) dan eror (ε) adalah independen, sehingga $E(F) = 0$ dan $Cov(F) = I$, I merupakan matriks identitas serta $E(\varepsilon) = 0$ dan $Cov(\varepsilon) = \Psi$, dengan Ψ merupakan diagonal matriks. Ψ didapatkan melalui persamaan:

$$\tilde{\Psi}_i = 1 - \tilde{h}_i^2 \quad (2.4)$$

$$h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2 \quad (2.5)$$

Keterangan:

h_i^2 : *communality*, dimana menunjukkan varians dari variabel bebas yang dapat menerangkan sejumlah k faktor atau komponen.

ℓ_{im}^2 : *loading factor* ke- i pada variabel ke- h

$\tilde{\Psi}_i$: proporsi varian darivariabel bebas yang disebabkan oleh faktor spesifik atau galat.

(Johnson, R. A., dan Wichern, D. D., 2007)

2.3 Jaringan Syaraf Tiruan

Jaringan saraf tiruan pertama kali diperkenalkan oleh McCulloch dan Pitts pada tahun 1943 (Siang, 2009). Jaringan syaraf tiruan didefinisikan sebagai suatu sistem pemrosesan informasi yang mempunyai karakteristik menyerupai jaringan saraf manusia. Jaringan saraf tiruan tercipta sebagai suatu generalisasi model matematis dari pemahaman manusia (*human cognition*) yang didasarkan pada asumsi sebagai berikut:

1. Pemrosesan informasi terjadi pada elemen sederhana yang disebut *neuron*.
2. Sinyal dikirimkan di antara *neuron-neuron* melalui penghubung.
3. Setiap penghubung memiliki bobot yang dapat menguatkan atau memperlemah sinyal.

4. Untuk menentukan *output*, setiap *neuron* menggunakan fungsi aktivasi, yang sebagian besar bukan merupakan fungsi linier, yang dikenakan pada jumlahan *input* yang diterima. Besarnya *output* ini kemudian dibandingkan dengan suatu nilai batas yang telah ditentukan.

Jaringan saraf tiruan mampu belajar dari pengalaman, melakukan generalisasi atas contoh-contoh yang telah diperoleh dan mengabstraksi karakteristik *input* bahkan untuk data yang tidak relevan. Berbeda dengan metode algoritma yang lain, jaringan syaraf tiruan beroperasi secara langsung dengan angka, sehingga data yang tidak numerik harus diubah ke dalam bentuk numerik. Jaringan saraf tiruan tidak memerlukan atau menggunakan suatu model matematis untuk menyelesaikan permasalahan yang dihadapi. Oleh karena itu jaringan saraf tiruan juga dikenal sebagai model *free estimator*.

Jaringan saraf tiruan memiliki sifat yang tidak transparan yaitu tidak menerangkan bagaimana suatu hasil didapatkan. Hal inilah yang membuat jaringan saraf tiruan mampu digunakan untuk menyelesaikan persoalan yang tidak terstruktur dan tidak dapat didefinisikan. Beberapa aplikasi jaringan saraf tiruan adalah:

1. Klasifikasi yaitu memilih suatu input data ke dalam suatu kategori tertentu yang ditetapkan.
2. Asosiasi yaitu menggambarkan suatu obyek secara keseluruhan hanya dengan sebuah bagian dari obyek lain.
3. *Self organizing* yaitu mampu mengolah data *input* tanpa memiliki data target.
4. Optimasi yaitu menemukan solusi yang paling baik sehingga seringkali mampu meminimalkan suatu fungsi biaya (*optimizer*).

Jaringan saraf tiruan memiliki kelemahan dalam melakukan operasi numerik dengan presisi yang tinggi, operasi algoritma aritmatik, operasi logika, operasi simbolis serta memerlukan waktu proses pelatihan yang cukup lama pada data dengan jumlah besar. Istilah yang digunakan pada jaringan saraf tiruan meru-

pakan istilah lain yang sering digunakan pada model statistika. Berikut adalah beberapa istilah pada jaringan saraf tiruan yang memiliki persamaan istilah pada model statistika:

Tabel 2.1 Istilah dalam Jaringan Saraf Tiruan

No.	Istilah dalam Statistika	Istilah dalam Jaringan Saraf Tiruan
1	Variabel	Fitur
2	Variabel independen	<i>Input</i>
3	Prediksi	<i>Output</i>
4	Variabel dependen	Target atau nilai <i>training</i>
5	Residual	<i>Errors</i>
6	Estimasi	<i>Training</i> , pembelajaran, adaptasi atau <i>self-organizing</i>
7	Kriteria estimasi	<i>Error function</i> , <i>cost function</i> atau <i>Lyapunov function</i>
8	Observasi	Pola atau <i>training pairs</i>
9	Estimasi parameter	Bobot
10	Interaksi	<i>High-order neurons</i>
11	Transformasi	<i>Function link</i>
12	Regresi atau analisis diskriminan	Pelatihan <i>supervised</i> atau <i>heteroassociation</i>
13	Reduksi data	Pelatihan <i>unsupervised</i> , <i>encoding</i> atau <i>autoassociation</i>
14	Analisis kluster	Pelatihan kompetisi atau <i>adaptive vector quantization</i>
15	Interpolasi dan ekstrapolasi	<i>Generalization</i>

(Sarle, 1994)

Pada jaringan saraf tiruan terdapat dua metode pelatihan, yaitu pelatihan *supervised* dan *unsupervised*. Metode pelatihan *supervised* mengasumsikan tersedianya target yang mengklasifikasikan contoh-contoh pelatihan ke dalam kelas-kelas. Algoritma pelatihan *supervised* memanfaatkan informasi dari setiap anggota kelas. Informasi tersebut bermanfaat untuk mendeteksi kesalahan klasifikasi. Pada metode pelatihan *unsupervised* tidak terdapat target, sehingga proses pengidentifikasian informasi kelas pada

pola merupakan proses dari pelatihan. Metode pelatihan *unsupervised* menggunakan contoh yang tidak diklasifikasikan jenisnya, sehingga sistem akan melakukan pemrosesan secara *heuristically*. Pelatihan *unsupervised* dapat dirancang untuk belajar dengan cepat dan lebih praktis dalam hal kecepatan.

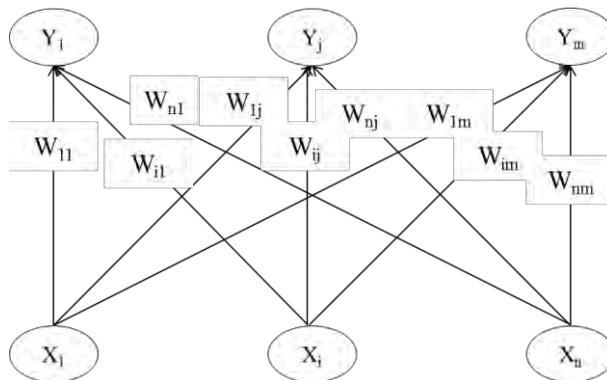
Seperti jaringan saraf manusia, jaringan saraf tiruan juga terdiri dari beberapa *neuron* dan pada setiap *neuron* terdapat hubungan. *Neuron* akan mentransformasikan informasi yang diterimanya melalui sambungan keluaran menuju *neuron-neuron* yang lain. Oleh karena itu dapat dikatakan bahwa *neuron* atau sel saraf adalah sebuah sel pemroses informasi yang merupakan dasar operasi dari jaringan syaraf tiruan. *Input* atau masukan yang menuju unit pengolah masing-masing berasal dari unit-unit yang berbeda x_n . Masing-masing sambungan memiliki kekuatan hubungan terkait (bobot) yang disimbolkan dengan w_{ij} . w_{ij} merupakan besarnya bobot yang menghubungkan antara unit ke x_i menuju ke unit y_j . Bobot yang diberikan pada setiap *input* didasarkan pada besarnya kekuatan hubungan. Kekuatan hubungan akan disesuaikan selama proses pelatihan sehingga pada akhir pelatihan akan dihasilkan bobot yang optimal. Unit pengolah akan membentuk penjumlahan berbobot dari setiap masukannya untuk menghitung hasil keluaran. Hasil perhitungan akan dikirimkan melalui hubungan keluaran.

(Fausset, L., 1994)

2.4 Kohonen SOM

Metode Kohonen SOM merupakan salah satu metode pelatihan *unsupervised*. Jaringan Kohonen SOM banyak digunakan untuk membagi pola masukan ke dalam beberapa kelompok (*cluster*). Dimisalkan masukan berupa vektor yang terdiri dari n komponen yang akan dikelompokkan ke dalam maksimum m buah kelompok yang disebut sebagai vektor contoh. Keluaran jaringan adalah kelompok yang paling dekat atau memiliki kemiripan dengan masukan yang diberikan. Ukuran kedekatan yang dipakai adalah jarak *Euclidean* yang paling minimum.

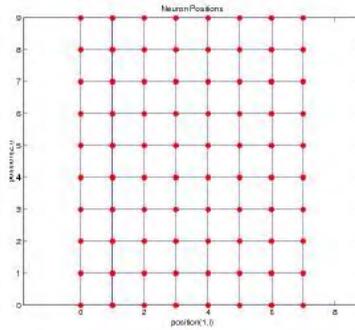
Bobot vektor-vektor contoh berfungsi sebagai penentu kedekatan vektor contoh tersebut dengan masukan yang diberikan. Selama masa pelatihan, vektor contoh yang paling dekat dengan masukan akan muncul sebagai pemenang. Pada vektor pemenang dan vektor-vektor yang berada di sekitarnya akan dilakukan penyesuaian bobot. Pada arsitektur jaringan Kohonen SOM tidak menggunakan perhitungan net, yaitu hasil kali vektor masukan dengan bobot dan juga fungsi aktivasi. Berikut ini adalah arsitektur jaringan Kohonen SOM :



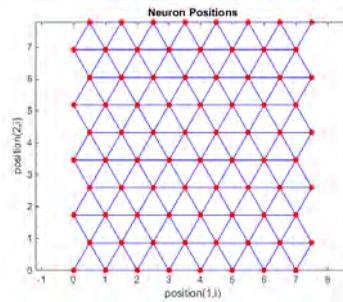
Gambar 2.1 Arsitektur Jaringan Kohonen SOM

Pada pelatihan *unsupervised* aturan Kohonen SOM digunakan pada beberapa jenis topologi jaringan. Diantaranya adalah jaringan kompetisi dan jaringan Kohonen SOM. Jika pada jaringan kompetisi neuron target diletakkan dalam sebuah baris (satu dimensi) sedangkan pada jaringan Kohonen SOM neuron target diletakkan dalam 2 dimensi yang bentuk/topologinya dapat diatur. Terdapat 3 macam jenis topologi dalam jaringan Kohonen SOM, yaitu sebagai berikut:

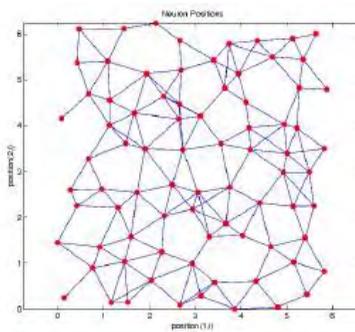
1. *Gridtop* yaitu neuron target disusun dalam array 2 dimensi
 2. *Hextop* yaitu neuron target disusun dalam bentuk heksagonal
 3. *Randtop* yaitu neuron disusun dalam 2 dimensi secara acak.
- Masing-masing jenis topologi *gridtop*, *hextop* dan *randtop* disajikan dengan gambar sebagai berikut:



Gambar 2.2 Topologi *Gridtop*



Gambar 2.3 Topologi *Hextop*



Gambar 2.4 Topologi *Randtop*

Algoritma jaringan kohonen untuk mengelompokan observasi adalah sebagai berikut:

0. Inisialisasi yang terdiri dari beberapa hal sebagai berikut:
 - a. Menetapkan inisialisasi bobot w_{ij}
 Nilai bobot w_{ij} ditentukan nilai yang dipilih secara *random* yang berada pada *range* nilai vektor *input*. Namun penentuan nilai bobot w_{ij} juga dapat dilakukan berdasarkan penelitian sebelumnya yang telah memperoleh hasil yang sesuai pada kasus penelitian yang sama.
 - b. Menetapkan nilai *learning rate* (α)
 Nilai *learning rate* yang digunakan merupakan nilai *default* (sudah ditetapkan) yaitu sebesar 0,02.
 - c. Menentukan nilai radius R
 Pada jaringan kohonen nilai radius R yang digunakan adalah $R = 0$. Ketika R bernilai 0 maka hanya vektor *input* pemenang yang akan melakukan proses pembelajaran.
1. Melakukan langkah 2 sampai 8 ketika kondisi untuk berhenti tidak terpenuhi
2. Melakukan langkah 3 sampai 5 untuk masing-masing vektor *input* x_i
3. Menghitung jarak Euclidean, yaitu jarak antara bobot w_{ij} dengan vektor *input* x_i untuk setiap kluster j menggunakan persamaan:

$$D_j = \sum_i (w_{ij} - x_i)^2 \quad (2.6)$$

Keterangan:

D_j = jarak *Euclidean*

w_{ij} = bobot yang menghubungkan antara vektor *input* x_i menuju ke unit y_j

4. Menentukan nilai J yang meminimumkan nilai D_j
5. Mengubah nilai bobot w_{ij} dari vektor *input* pemenang. Bobot w_{ij} yang baru ditulis menggunakan persamaan:

$$w_{ij} (\text{new}) = w_{ij} (\text{old}) + \alpha [x_i - w_{ij} (\text{old})] \quad (2.7)$$

Keterangan:

$w_{ij}(\text{new})$ = bobot w_{ij} yang baru

$w_{ij}(\text{old})$ = bobot w_{ij} awal

α = nilai *learning rate*

x_i = vektor *input* ke- i

6. Memperbarui nilai *learning rate* menggunakan persamaan:

$$\alpha(t+1) = 0,5x\alpha(t) \quad (2.8)$$

7. Mereduksi nilai radius R untuk mencapai nilai minimum *learning rate* yang diinginkan.
8. Menguji kondisi penghentian iterasi.

Kohonen menyebutkan bahwa pengurangan nilai *learning rate* secara linier dan geometrik akan memberikan hasil yang cukup baik pada saat melakukan proses pengulangan (iterasi). Nilai radius R, yang merupakan radius kedekatan di sekitar kluster, akan mengalami reduksi selama proses pengulangan (iterasi) berlangsung. Proses pembentukan kluster didasarkan pada dua hal, yaitu jumlah pengulangan yang telah ditetapkan dan kondisi konvergen hasil akhir. Pencapaian kondisi konvergen pada hasil akhir membutuhkan waktu yang cukup lama dan nilai *learning rate* yang kecil. Berdasarkan hal tersebut seringkali pengulangan (iterasi) didasarkan pada suatu nilai yang telah ditetapkan sebelumnya.

(Fausset, L., 1994)

2.5 Metode *K-Means*

K-Means Cluster Analysis merupakan salah satu metode analisis kluster non hirarki yang dapat digunakan untuk mempartisi objek ke dalam kelompok-kelompok berdasarkan kedekatan karakteristik, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu kluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kluster yang lain. Tujuan pengelompokan adalah untuk meminimalkan *objective function* yang di set dalam proses pengelompokan, yang pada dasarnya berusaha untuk meminimalkan

variasi dalam satu kluster dan memaksimalkan variasi antar kluster.

Jika diberikan sekumpulan objek $X = X_1, X_2, \dots, X_n$ maka algoritma *K-Means Cluster Analysis* akan mempartisi X dalam k buah kluster, setiap kluster memiliki *centroid* dari objek-objek dalam kluster tersebut. Pada tahap awal algoritma *K-Means Cluster Analysis* dipilih secara acak k buah objek sebagai *centro-id*, kemudian jarak antara objek dengan *centroid* dihitung dengan menggunakan jarak *Euclidian*, objek ditempatkan dalam kluster yang terdekat dihitung dari titik tengah kluster. *Centroid* baru ditetapkan jika semua objek sudah ditempatkan dalam kluster terdekat. Proses penentuan *centroid* dan penempatan objek dalam kluster diulangi sampai nilai *centroid* konvergen (*centroid* dari semua kluster tidak berubah lagi). Secara umum metode *K-Means Cluster Analysis* menggunakan algoritma sebagai berikut:

1. Menentukan K sebagai jumlah kluster yang dibentuk.

Untuk menentukan banyaknya kluster K dilakukan dengan beberapa pertimbangan seperti pertimbangan teoritis dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak kluster.

2. Membangkitkan K *centroid* (titik pusat kluster) awal secara *random*.

Penentuan *centroid* awal dilakukan secara *random* dari objek-objek yang tersedia sebanyak K kluster, kemudian untuk menghitung *centroid* kluster ke- i berikutnya, digunakan rumus sebagai berikut:

$$y_i = \frac{\sum_{i=1}^n x_i}{n} \quad (2.9)$$

Keterangan:

y_i = *centroid* pada *cluster*

x_i = obyek pengamatan ke i

n = banyaknya obyek yang menjadi anggota kluster.

3. Menghitung jarak setiap objek ke masing-masing *centroid* pada setiap kluster.

Untuk menghitung jarak antara objek dengan *centroid* menggunakan *Euclidian Distance* dengan persamaan sebagai berikut:

$$d = \sum_{i=1}^n (x_i - y_i)^2 \quad (2.10)$$

Keterangan:

x_i = obyek pengamatan ke i

y_i = *centroid* ke i

n = banyaknya obyek yang menjadi anggota kluster

4. Mengalokasikan masing-masing objek ke dalam *centroid* yang paling terdekat.

Setiap objek dinyatakan sebagai anggota kluster dengan mengukur jarak kedekatan sifatnya terhadap titik pusat kluster tersebut.

5. Melakukan iterasi, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan (2.9)

6. Mengulangi langkah 3 jika posisi *centroid* baru tidak sama.

Pengecekan *konvergensi* dilakukan dengan membandingkan matriks *group assignment* pada iterasi sebelumnya dengan matrik *group assignment* pada iterasi yang sedang berjalan. Jika hasilnya sama maka algoritma *K-Means* sudah *konvergen*, tetapi jika berbeda maka belum *konvergen* sehingga perlu dilakukan iterasi berikutnya.

(Johnson dan Wichern, 2007)

2.6 Pemilihan Jumlah Kluster Optimum

Untuk mengetahui jumlah kelompok optimum dapat digunakan kriteria nilai *Pseudo F-statistics*. Berikut adalah persamaan yang digunakan untuk menghitung nilai *Pseudo F-statistics*:

$$Pseudo\ Fstatistics = \frac{\left(\frac{R^2}{k-1} \right)}{\left(\frac{1-R^2}{n-k} \right)} \quad (2.11)$$

dimana:

$$R^2 = \frac{(SST - SSW)}{SST} \quad (2.12)$$

$$SST = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (x_{ijk} - \bar{x}_j)^2 \quad (2.13)$$

$$SSW = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (x_{ijk} - \bar{x}_{jk})^2 \quad (2.14)$$

Keterangan:

SST (*Sum Square Total*) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata keseluruhan

SSW (*Sum Square Within*) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata kelompoknya

n : banyaknya sampel

c : banyaknya variabel

p : banyaknya kelompok

x_{ijk} : sampel ke- i pada variabel ke- j kelompok ke- k

\bar{x}_j : rata-rata seluruh sampel pada variabel ke- j

\bar{x}_{jk} : rata-rata sampel pada variabel ke- j dan kelompok ke- k

Nilai *Pseudo F-statistics* tertinggi menunjukkan bahwa jumlah kelompok yang digunakan untuk memartisi data telah optimal. Dimana keragaman dalam kelompok sangat homogen sedangkan antar kelompok sangat heterogen.

(Calinski dan Harabasz, 1974 dalam Orpin dan Kostylev, 2006)

2.7 Kriteria Perbandingan Kluster Terbaik

Kluster terbaik untuk mengelompokkan obyek memiliki nilai homogenitas antar anggota dalam satu kluster yang tinggi dan memiliki nilai heterogenitas antar kluster yang tinggi. Dengan kata lain variasi antar anggota dalam satu kluster harus rendah sedangkan variasi antar kluster harus tinggi. Untuk menentukan kriteria kluster terbaik dapat digunakan nilai *icdrate*. *Icdrate* (*internal cluster dispersion*) menggambarkan tingkat dispersi atau perbedaan dalam kluster. Nilai *icdrate* yang semakin kecil me-

nunjukkan bahwa kluster tersebut semakin baik sebab antara anggota dalam satu kluster memiliki perbedaan yang rendah atau memiliki variasi yang kecil.

Nilai *icdrate* dapat diperoleh dengan menggunakan persamaan:

$$icdrate = 1 - \frac{SSB}{SST} \quad (2.15)$$

Dimana:

$$SSB = \sum_{j=1}^c \sum_{k=1}^p (\bar{x}_{jk} - \bar{x}_j)^2 \quad (2.16)$$

Keterangan:

SST (*Sum Square Total*) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata keseluruhan

\bar{x}_{jk} : rata-rata sampel pada variabel ke-*j* dan kelompok ke-*k*

\bar{x}_j : rata-rata sampel pada variabel ke-*j*

c : banyaknya variabel

p : banyaknya kelompok

(Mingoti dan Lima, 2006)

2.8 Pengujian *One-way* MANOVA

Pengujian *One-way* MANOVA digunakan untuk membandingkan nilai vektor *mean* antar perlakuan pada data multivariat. Dimana pada pengujian *One-way* MANOVA hanya menggunakan satu faktor atau perlakuan dan tanpa mempertimbangkan interaksi antar perlakuan. Berikut adalah hipotesis yang digunakan dalam pengujian *One-way* MANOVA:

$H_0: \tau_1 = \tau_2 = \tau_3 = \dots = \tau_g = 0$

$H_1: \text{minimal ada 1 } \tau_l \neq 0, \text{ dengan } l = 1, 2, \dots, g$

Statistik uji:

$$F = \frac{SS_{\text{treatment}} / (g - 1)}{SS_{\text{residual}} / \left(\sum_{l=1}^g n_l - g \right)} \quad (2.17)$$

Keterangan:

$SS_{treatment}$ = nilai *Sum of Square* perlakuan

$SS_{residual}$ = nilai *Sum of Square* residual

n_l = banyaknya data populasi ke l

g = banyaknya populasi

Tolak H_0 jika $F > F_{v_1, v_2, \alpha}$, dengan $v_1 = g - 1$ dan $v_2 = \sum n_l - g$.

Berikut ini adalah tabel uji *One-way* MANOVA:

Tabel 2.2 Tabel Uji *One-way* MANOVA

Sumber Variasi	Matrik Jumlah Kuadrat	Derajat Bebas (db)
Perlakuan	$B = \sum_{l=1}^g n_l (\bar{X}_l - \bar{X})(\bar{X}_l - \bar{X})'$	$g-1$
Residual (Error)	$W = \sum_{l=1}^g \sum_{j=1}^{n_l} (\bar{X}_{lj} - \bar{X}_l)(\bar{X}_{lj} - \bar{X}_l)'$	$\sum_{l=1}^g n_l - g$
Total	$B + W = \sum_{l=1}^g \sum_{j=1}^{n_l} (X_{lj} - \bar{X})(X_{lj} - \bar{X})'$	$\sum_{l=1}^g n_l - 1$

(Johnson dan Winchern, 2007)

2.9 Indeks Kesehatan Masyarakat

Indikator kesehatan dapat digunakan untuk mengetahui kondisi kesehatan masyarakat di suatu wilayah. Dimana terdapat beberapa pendapat tentang variabel yang membentuk Indikator Kesehatan Masyarakat. Menurut Indeks Pembangunan Kesehatan Manusia (IPKM), terdapat 24 indikator kesehatan yang digunakan dalam sebagai ukuran tingkat pembangunan kesehatan manusia disuatu wilayah. Indikator kesehatan tersebut adalah prevalensi balita gizi buruk, prevalensi balita sangat pendek dan pendek, prevalensi balita sangat kurus dan kurus, prevalensi balita gemuk, prevalensi diare, prevalensi pnemonia, prevalensi hipertensi, prevalensi gangguan mental, prevalensi asma, prevalensi penyakit gigi dan mulut, prevalensi disabilitas, prevalensi cedera, prevalensi penyakit sendi, prevalensi ISPA, proporsi perilaku cuci tangan, proporsi

merokok tiap hari, akses air bersih, akses sanitasi, cakupan persalinan oleh nakes, cakupan pemeriksaan neonatal-1, cakupan imunisasi lengkap, cakupan penimbangan balita, rasio Dokter/Puskesmas, dan rasio bidan/desa.

Menurut *World Health Organization*, indikator kesehatan masyarakat didasarkan pada hal-hal sebagai berikut:

1. Berhubungan dengan status kesehatan masyarakat, yang terdiri dari:
 - a. Indikator komprehensif, yaitu angka kematian kasar menurun, rasio angka mortalitas proporsional rendah dan umur harapan hidup meningkat.
 - b. Indikator spesifik, yaitu angka kematian ibu dan anak menurun, angka kematian karena penyakit menular menurun dan angka Angka kelahiran menurun.
2. Berhubungan dengan pelayanan kesehatan, yang terdiri dari:
 - a. Rasio antara pelayanan kesehatan dan jumlah penduduk seimbang.
 - b. Distribusi tenaga kesehatan merata.
 - c. Informasi lengkap tentang fasilitas kesehatan.
 - d. Informasi tentang sarana Pelayanan kesehatan di Rumah Sakit, Puskesmas, dan fasilitas pelayanan kesehatan yang lain.

Departemen Kesehatan Republik Indonesia menyebutkan bahwa indikator kesehatan masyarakat menurut Indonesia Sehat 2010 terdiri dari tiga indikator, yaitu:

1. Indikator Hasil Akhir, terdiri atas indikator angka-angka mortalitas, angka-angka morbiditas, dan indikator status gizi.
2. Indikator Hasil Antara, terdiri atas indikator keadaan lingkungan, indikator perilaku hidup masyarakat, dan indikator akses dan mutu pelayanan kesehatan.
3. Indikator Proses dan Masukan, terdiri atas indikator pelayanan kesehatan, indikator sumber daya kesehatan, indikator manajemen kesehatan serta indikator kontribusi sektor-sektor terkait.

Gerring et al (2013) melakukan penelitian untuk mengukur kinerja sistem kesehatan berdasarkan Indikator Kesehatan Masya-

rakat. Indikator Kesehatan Masyarakat dipengaruhi oleh beberapa variabel diantaranya variabel ekonomi, pendidikan dan epidemiologi. Namun bukan berarti Indikator Kesehatan Masyarakat ditentukan berdasarkan ketiga variabel tersebut. Variabel ekonomi yang membentuk Indikator Kesehatan Masyarakat terdiri dari nilai PDRB (Pendapatan Domestik Regional Bruto) per kapita pada masing-masing daerah dan Angka Harapan Hidup. Secara teori disebutkan bahwa apabila suatu daerah memiliki PDRB per kapita dan Angka Harapan Hidup yang tinggi atau dapat dikatakan bahwa daerah tersebut merupakan daerah yang memiliki kemampuan ekonomi yang baik, maka diharapkan daerah tersebut juga mempunyai pengeluaran untuk sektor kesehatan yang tinggi dan baik pula. Oleh sebab itu dapat dikatakan bahwa terdapat hubungan antara kondisi kesehatan di suatu daerah dengan kondisi ekonomi di daerah tersebut.

Variabel pendidikan dapat diukur berdasarkan nilai Angka Melek Huruf dan rata-rata total lama sekolah. Dimana suatu daerah dengan kondisi pendidikan yang baik diharapkan dapat memberikan kontribusi bagi peningkatan kondisi ekonomi di wilayah tersebut. Variabel epidemiologi mencerminkan penyakit menular yang disebabkan oleh vektor. Variabel epidemiologi yang digunakan merupakan angka prevalensi penyakit TBC dan malaria. Prevalensi penyakit menular melalui vektor di suatu daerah diduga dipengaruhi oleh angka prevalensi penyakit menular di daerah yang letaknya dekat dengan daerah tersebut.

(Halaman ini sengaja dikosongkan)

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan pada penelitian ini merupakan data sekunder tentang variabel pembentuk Indikator Kesehatan Masyarakat di 38 kabupaten/kota yang diperoleh dari hasil Riskesdas dan Badan Pusat Statistik Provinsi Jawa Timur pada tahun 2012.

3.2 Variabel Penelitian

Variabel penelitian yang digunakan merupakan variabel pembentuk Indikator Kesehatan Masyarakat, yang meliputi variabel ekonomi, pendidikan dan epidemiologi dengan rincian sebagai berikut:

Tabel 3.1 Variabel Penelitian

Variabel	Nama Variabel
X_1	PRDB (Pendapatan Regional Domestik Bruto) per kapita atas dasar harga konstan
X_2	Angka Harapan Hidup
X_3	Angka Melek Huruf
X_4	Rata-rata total lama sekolah penduduk usia 15 tahun
X_5	Prevalensi TBC (Tuberkulosis)
X_6	Prevalensi malaria

Berikut ini adalah penjelasan masing-masing variabel penelitian yang digunakan:

- a. Variabel ekonomi terdiri dari PRDB (Pendapatan Regional Domestik Bruto) per kapita atas dasar harga konstan dan Angka Harapan Hidup.
 1. Variabel X_1 yaitu PRDB (Pendapatan Regional Domestik Bruto) per kapita atas dasar harga konstan menyatakan nilai keseluruhan semua barang dan jasa yang diproduksi dalam suatu wilayah dalam suatu jangka waktu tertentu.
 2. Variabel X_2 yaitu Angka Harapan Hidup menyatakan rata-rata tahun hidup yang masih akan dijalani oleh seseorang

yang telah berhasil mencapai umur x , pada suatu tahun tertentu, dalam situasi mortalitas yang berlaku di lingkungan masyarakatnya.

- b. Variabel pendidikan terdiri dari Angka Melek Huruf dan Rata-rata total lama sekolah penduduk usia 15 tahun.
 1. Variabel X_3 yaitu Angka Melek Huruf menyatakan persentase penduduk usia 15 tahun yang bisa membaca dan menulis serta mengerti sebuah kalimat sederhana dalam hidupnya sehari-hari.
 2. Variabel X_4 yaitu rata-rata total lama sekolah penduduk usia 15 tahun, menyatakan rata-rata lama waktu untuk lama sekolah yang dilakukan oleh penduduk dengan kategori usia 15 tahun.
- c. Variabel epidemiologi yaitu variabel yang terkait dengan wabah penyakit yang terjadi di masyarakat yang terdiri dari prevalensi TBC dan prevalensi malaria.
 1. Variabel X_5 yaitu prevalensi TBC menyatakan jumlah orang dalam populasi yang menderita penyakit TBC pada waktu tertentu.
 2. Variabel X_6 yaitu prevalensi malaria menyatakan jumlah orang dalam populasi yang menderita penyakit malaria pada waktu tertentu.

3.3 Langkah Analisis

Langkah analisis awal yang dilakukan adalah mendeskripsikan karakteristik masing-masing Indikator Kesehatan Masyarakat berdasarkan kabupaten/kota di Jawa Timur di Indonesia dengan menggunakan nilai rata-rata, varian, nilai maksimum dan nilai minimum. Langkah selanjutnya adalah melakukan pengelompokan kabupaten/kota di Jawa Timur menggunakan Kohonen SOM dan *K-Means*. Sebelum melakukan pengelompokan terlebih dahulu dilakukan Uji Bartlett untuk mengetahui ada atau tidaknya hubungan antara variabel penelitian. Apabila pada hasil pengujian diperoleh kesimpulan bahwa terdapat hubungan antar variabel penelitian, maka selanjutnya dilakukan analisis faktor untuk me-

reduksi variabel yang saling berkorelasi. Sebaliknya, apabila tidak ditemukan korelasi antar variabel penelitian maka langsung dapat dilakukan pengelompokan dapat dilakukan.

Pada masing-masing metode Kohonen SOM dan *K-Means* akan digunakan jumlah kelompok sebanyak 2 klaster dan 3 klaster. Penentuan jumlah klaster tersebut dilakukan berdasarkan penelitian sebelumnya yang telah dilakukan oleh Ramadhani (2014). Kemudian ditentukan jumlah kelompok optimum berdasarkan nilai *Pseudo F-statistics*. Kemudian dilakukan pengelompokan menggunakan metode Kohonen SOM dan *K-Means* yang didasarkan pada jumlah kelompok optimum. Langkah yang terakhir kelompok yang terbentuk dibandingkan berdasarkan nilai *icdrate* untuk menentukan hasil pengelompokan terbaik serta melakukan pengujian *One-way* MANOVA yang bertujuan untuk mengetahui apakah masing-masing variabel yang membentuk kelompok memiliki perbedaan.

Rincian tahapan analisis adalah sebagai berikut:

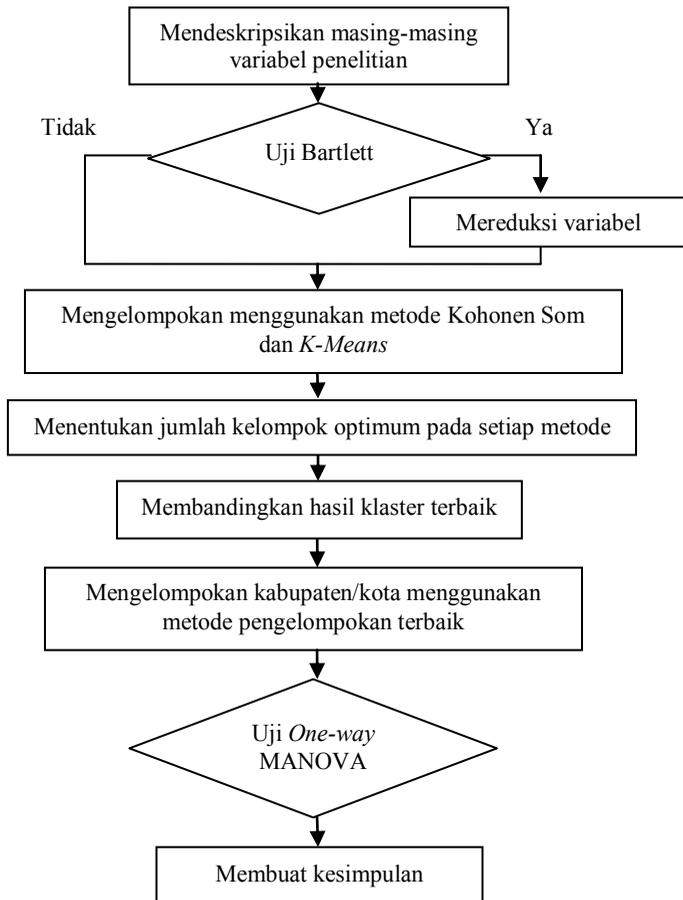
1. Mendeskripsikan karakteristik masing-masing indikator kesehatan masyarakat berdasarkan kabupaten/kota di Jawa Timur.
2. Melakukan Uji Bartlett. Apabila pada hasil pengujian diperoleh kesimpulan bahwa terdapat hubungan antar variabel penelitian, maka selanjutnya dilakukan analisis faktor untuk mereduksi variabel yang saling berkorelasi. Sebaliknya, apabila tidak ditemukan korelasi antar variabel penelitian maka langsung dapat dilakukan pengelompokan dapat dilakukan.
3. Mengelompokan kabupaten/kota di Jawa Timur menggunakan metode Kohonen SOM untuk masing-masing ukuran topologi menjadi 2 klaster dan 3 klaster, yang meliputi tahapan sebagai berikut:
 - a. Menetapkan inisialisasi bobot w_{ij} , menetapkan nilai *learning rate* (α) dan menentukan nilai radius R.
 - b. Menghitung jarak Euclidean, yaitu jarak antara bobot w_{ij} dengan vektor contoh x_i untuk setiap klaster j dengan menggunakan persamaan (2.6)
 - c. Menentukan nilai J yang meminimumkan nilai D_j

- d. Mengubah nilai bobot w_{ij} pada vektor *input* pemenang menggunakan persamaan (2.7)
- e. Memperbarui nilai *learning rate* yang dapat dihitung menggunakan persamaan (2.8)
- f. Mereduksi nilai radius R untuk mencapai nilai minimum *learning rate* yang diinginkan.
- g. Menguji kondisi penghentian iterasi.
4. Melakukan pengelompokan menggunakan metode *K-Means*, yang memiliki tahapan sebagai berikut
 - a. Menentukan jumlah kluster yang dibentuk.
 - b. Menentukan pusat kluster secara *random*. Nilai *centroid* dapat dihitung menggunakan persamaan (2.9)
 - c. Menghitung jarak setiap objek ke masing-masing *centroid* dari masing-masing kluster menggunakan persamaan (2.10)
 - d. Mengalokasikan masing-masing objek ke dalam *centroid* yang paling terdekat.
 - e. Melakukan iterasi.
 - f. Mengulangi langkah 3 jika posisi *centroid* baru tidak sama.
5. Mendapatkan jumlah kelompok optimum untuk masing-masing metode pengelompokan berdasarkan nilai *Pseudo F-statistics* pada persamaan (2.11).
6. Memilih hasil kelompok terbaik berdasarkan metode Kohonen SOM dan *K-Means* dengan menggunakan kriteria nilai *icdrate* pada persamaan (2.15) dengan tahapan sebagai berikut:
 - a. Menghitung nilai *Sum Square Total* (SST) dengan menggunakan persamaan (2.13)
 - b. Menghitung nilai *Sum Square Between* (SSB) menggunakan persamaan (2.16)
 - c. Menghitung nilai *icdrate* dengan menggunakan persamaan (2.15)
 - d. Membandingkan nilai *icdrate* untuk memperoleh hasil pengelompokan terbaik.
7. Mengelompokan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat menggunakan metode pengelompokan terbaik. Setelah kelompok terbentuk kemudian

dilakukan pengujian *One-way* MANOVA untuk membuktikan apakah masing-masing variabel penelitian yang digunakan untuk membentuk kelompok memiliki perbedaan di masing-masing kluster.

8. Membuat kesimpulan.

Secara ringkas langkah analisis disajikan dalam diagram alir sebagai berikut:



Gambar 3.1 Diagram Alir Penelitian

(Halaman ini sengaja dikosongkan)

BAB IV

ANALISIS DAN PEMBAHASAN

Pada bab berikut ini akan dilakukan analisis dan pembahasan tentang indikator kesehatan masyarakat dan variabel yang membentuknya di Jawa Timur pada tahun 2012. Analisis dan pembahasan yang akan dilakukan terdiri dari deskripsi data tentang variabel yang membentuk Indikator Kesehatan Masyarakat kemudian melakukan pengelompokan kabupaten/kota di Jawa Timur. Sebelum dilakukan pengelompokan, terlebih dahulu dilakukan Uji Bartlett untuk mengetahui ada atau tidaknya korelasi diantara variabel penelitian. Apabila terdapat korelasi maka selanjutnya dilakukan reduksi variabel menggunakan analisis faktor, sedangkan apabila tidak terdapat korelasi maka pengelompokan dapat langsung dilakukan.

Pengelompokan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat terdiri dari beberapa tahapan, yaitu membentuk kabupaten/kota di Jawa Timur menjadi 2 klaster dan 3 klaster untuk masing-masing metode *K-Means* dan metode Kohonen SOM dengan tiga jenis topologi. Selanjutnya dari beberapa kelompok yang terbentuk tersebut dibandingkan berdasarkan nilai *Pseudo F-statistics* untuk memperoleh jumlah kelompok optimum. Selain itu hasil kelompok yang terbentuk juga dibandingkan berdasarkan nilai *icdrate* untuk menentukan hasil pengelompokan terbaik.

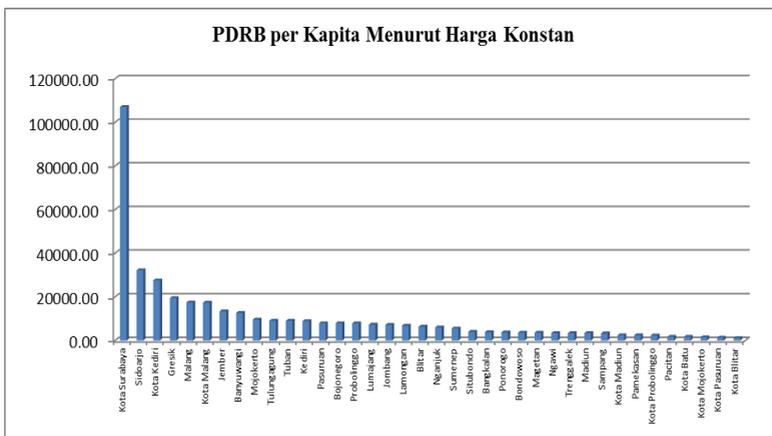
4.1 Statistika Deskriptif

Sebelum dilakukan analisis untuk menentukan jumlah kelompok optimum pada masing-masing metode Kohonen SOM dan *K-Means*, terlebih dahulu akan dilakukan deskripsi data tentang variabel yang membentuk Indikator Kesehatan Masyarakat di Jawa Timur pada tahun 2012. Untuk melakukan deskripsi data akan digunakan beberapa ukuran pemusatan data yang terdiri dari nilai rata-rata, varian, nilai maksimum dan nilai minimum. Selain itu data juga akan disajikan dalam bentuk grafik. Pembahasan

statistika deskriptif akan dilakukan berdasarkan variabel yang membentuk Indikator Kesehatan Masyarakat di Jawa Timur.

4.1.1 Statistika Deskriptif Variabel Ekonomi

Variabel ekonomi yang membentuk Indikator Kesehatan Masyarakat di Jawa Timur terdiri dari variabel Produk Domestik Bruto (PDRB) per kapita menurut harga konstan dan Angka Harapan Hidup. Berikut ini adalah nilai PDRB per kapita daerah-daerah di Jawa Timur dalam bentuk grafik:



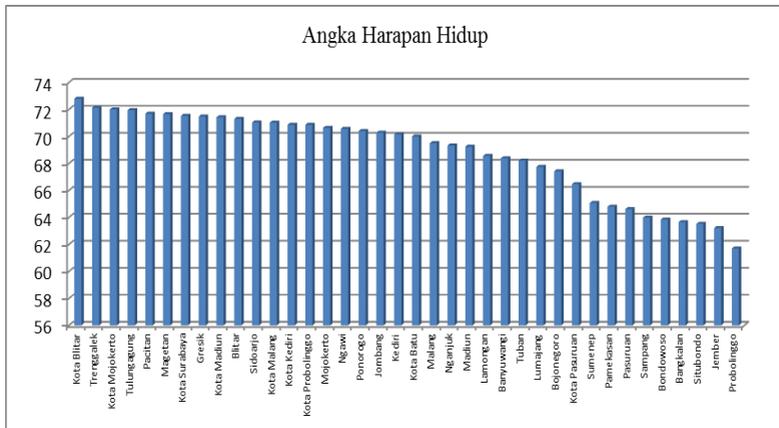
Gambar 4.1 Grafik PDRB per Kapita Menurut Harga Konstan

Produk Domestik Regional Bruto (PDRB) per kapita atas dasar harga konstan berguna untuk mengetahui pertumbuhan nyata ekonomi per kapita penduduk di suatu daerah. Hasil deskripsi data yang tertera pada Gambar 4.1 terlihat bahwa Kota Surabaya merupakan daerah yang memiliki nilai PDRB per kapita tertinggi di Provinsi Jawa Timur yaitu sebesar 107.059,48 (milyar rupiah). Di lain sisi nilai PDRB per kapita terendah yaitu sebesar 1.155,93 (milyar rupiah) dimiliki oleh Kota Blitar. Secara visual pada Gambar 4.1 juga terlihat bahwa antara daerah dengan PDRB per kapita tertinggi dan yang terendah dan yang tertinggi memiliki rentang nilai yang cukup jauh.

Kabupaten dan kota di Jawa Timur memiliki nilai rata-rata PDRB per kapita atas dasar harga konstan sebesar 10.360 (milyar rupiah). Angka tersebut merupakan angka yang cukup besar, namun hal tersebut tidak dapat digunakan untuk menyimpulkan bahwa pertumbuhan ekonomi di Jawa Timur telah baik sebab nilai varian PDRB per kapita atas harga konstan di Jawa Timur adalah sebesar 309.001.810. Nilai varian yang sangat besar tersebut mengindikasikan bahwa pertumbuhan ekonomi di Jawa Timur masih belum merata pada masing-masing kabupaten/kota. Hal tersebut juga didukung dengan hasil deskripsi data yang menunjukkan bahwa nilai minimum dan maksimum PDRB per kapita memiliki rentang yang juga cukup besar. Produk Domestik Bruto Regional (PDRB) merupakan nilai keseluruhan produk dan jasa yang diproduksi oleh wilayah tertentu dalam suatu jangka waktu tertentu. Sehingga tidak meratanya PDRB per kapita di Jawa Timur juga disebabkan oleh masing-masing daerah di Jawa Timur memiliki potensi sumber daya alam dan manusia yang berbeda-beda sesuai letak geografis, kondisi alam dan lingkungan.

Variabel ekonomi kedua yang membentuk Indikator Kesejahteraan Masyarakat di Jawa Timur adalah Angka Harapan Hidup. Angka Harapan Hidup merupakan alat untuk mengevaluasi kinerja pemerintah dalam meningkatkan kesejahteraan penduduk pada umumnya dan secara khusus untuk meningkatkan derajat kesehatan. Diketahui bahwa nilai rata-rata Angka Harapan Hidup di Jawa Timur pada tahun 2012 adalah sebesar 68,76 tahun. Nilai rata-rata tersebut masih berada di bawah rata-rata Angka Harapan Hidup nasional yaitu sebesar 69,87 tahun. Selain itu berdasarkan informasi pada Gambar 4.2, wilayah di Jawa Timur yang memiliki Angka Harapan Hidup terendah adalah Kabupaten Probolinggo yaitu sebesar 61,7 tahun dan Angka Harapan Hidup tertinggi dimiliki oleh Kota Blitar yaitu sebesar 72,8 tahun. Rentang nilai minimum dan maksimum pada Angka Harapan Hidup di Jawa Timur relatif kecil. Hal tersebut juga didukung dengan deskripsi data bahwa nilai varian pada Angka Harapan Hidup di Jawa Timur sebesar 9,885 sehingga dapat dikatakan bahwa Angka

Harapan Hidup di Jawa Timur telah merata pada masing-masing kabupaten/kota meskipun masih berada di bawah rata-rata Angka Harapan Hidup nasional. Angka Harapan Hidup di Jawa Timur yang masih berada di bawah rata-rata Angka Harapan Hidup nasional memerlukan penanganan khusus, yang meliputi program pembangunan kesehatan dan program sosial lainnya yang terdiri dari pembangunan kesehatan lingkungan, kecukupan gizi dan kalori serta program pemberantasan kemiskinan. Berikut ini adalah Gambar 4.2 yang menggambarkan nilai Angka Harapan Hidup pada masing-masing kabupaten/kota di Jawa Timur:

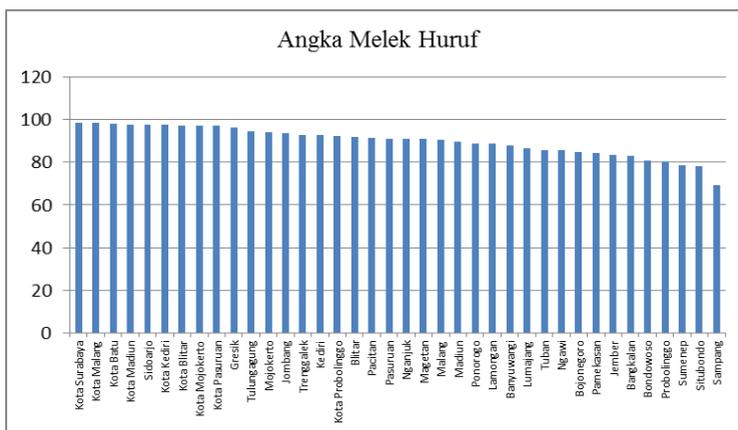


Gambar 4.2 Grafik Angka Harapan Hidup di Jawa Timur

4.1.2 Statistika Deskriptif Variabel Pendidikan

Variabel pendidikan yang membentuk Indikator Kesehatan Masyarakat di Jawa Timur adalah Angka Melek Huruf dan rata-rata lama sekolah. Angka Melek Huruf mencerminkan potensi perkembangan intelektual suatu daerah dan juga menunjukkan kontribusi terhadap pembangunan daerah. Pada tahun 2012 Jawa Timur memiliki rata-rata Angka Melek Huruf sebesar 90,03 yang masih berada di bawah rata-rata Angka Melek Huruf nasional di tahun 2012 yaitu sebesar 93,25.

Berdasarkan grafik pada Gambar 4.3 terlihat bahwa Angka Melek Huruf terbesar di Jawa Timur dimiliki oleh Kota Surabaya yaitu sebesar 98,35 dan Angka Melek Huruf terendah dimiliki oleh Kabupaten Sampang yaitu sebesar 69,12. Di Jawa Timur terdapat kabupaten/kota yang memiliki Angka Melek Huruf yang berada di atas rata-rata Angka Melek Huruf nasional, yaitu Kota Surabaya, namun juga terdapat kabupaten/kota, yaitu Kabupaten Sampang, yang berada jauh di bawah rata-rata Angka Melek Huruf nasional. Oleh sebab itu dapat dikatakan bahwa Angka Melek Huruf di Jawa Timur masih belum merata di masing-masing kabupaten/kota. Berikut ini adalah Gambar 4.3 yang mempresentasikan Angka Melek Huruf di Provinsi Jawa Timur pada setiap kabupaten/kota:

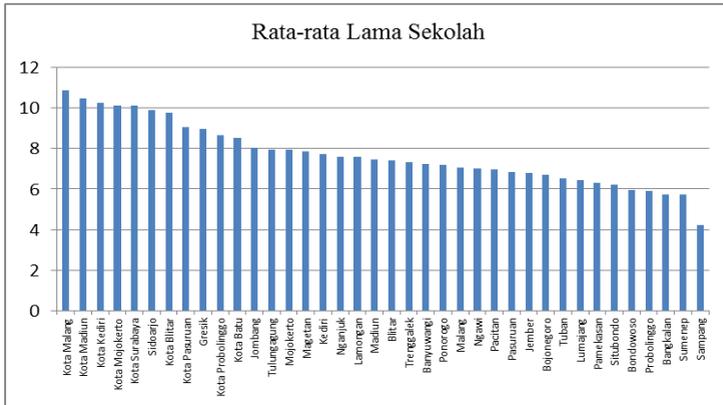


Gambar 4.3 Grafik Angka Melek Huruf di Jawa Timur

Rata-rata lama sekolah merupakan ukuran akumulasi investasi pendidikan individu sehingga diharapkan dapat dijadikan sebagai ukuran akumulasi modal manusia suatu daerah. Nilai rata-rata lama sekolah terbesar di Jawa Timur dimiliki oleh Kota Malang yaitu sebesar 10,87 tahun sedangkan nilai rata-rata terendah dimiliki oleh Kabupaten Sampang yaitu sebesar 4,22 tahun. Berdasarkan informasi sebelumnya diketahui bahwa Kabupaten

Sampang juga memiliki Angka Melek Huruf terendah di Jawa Timur, sehingga dapat dikatakan bahwa rata-rata lama sekolah di Kabupaten Sampang yang rendah mengakibatkan Angka Melek Huruf di wilayah tersebut juga rendah.

Rata-rata lama sekolah pada setiap kabupaten/kota di Jawa Timur disajikan pada Gambar 4.4 sebagai berikut:



Gambar 4.4 Grafik Rata-rata Lama Sekolah di Jawa Timur

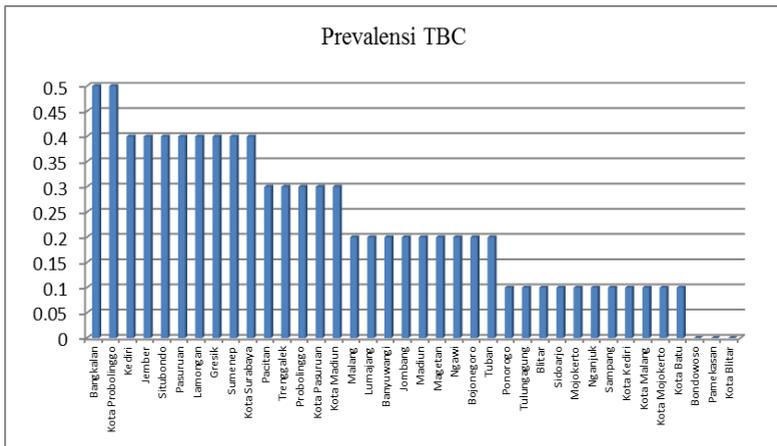
Pemerintah Indonesia mencanangkan program Wajib Belajar selama 9 tahun. Namun berdasarkan Gambar 4.2 terlihat bahwa masih terdapat kabupaten/kota di Jawa Timur yang memiliki nilai rata-rata lama sekolah yang kurang dari 9 tahun. Hal tersebut mengindikasikan bahwa masih banyak daerah di Jawa Timur yang memiliki tingkat pendidikan yang rendah. Berdasarkan Gambar 4.2 diketahui bahwa daerah yang memiliki rata-rata lama sekolah di atas 9 tahun adalah Kabupaten Pasuruan, Sidoarjo, Kota Surabaya, Kota Mojokerto, Kota Kediri, Kota Madiun dan Kota Malang.

4.1.3 Statistika Deskriptif Variabel Epidemiologi

Prevalensi penyakit menyatakan jumlah penduduk yang menderita suatu penyakit di wilayah tertentu dalam kurun tertentu dibagi dengan jumlah penduduk di wilayah tersebut. Penyakit

TBC dan malaria merupakan jenis penyakit menular. Penyakit TBC merupakan penyakit yang dapat ditularkan melalui udara sedangkan penyakit malaria adalah penyakit menular melalui vektor.

Berikut adalah grafik yang menunjukkan tingkat prevalensi TBC di Provinsi Jawa Timur di masing-masing kabupaten/kota:

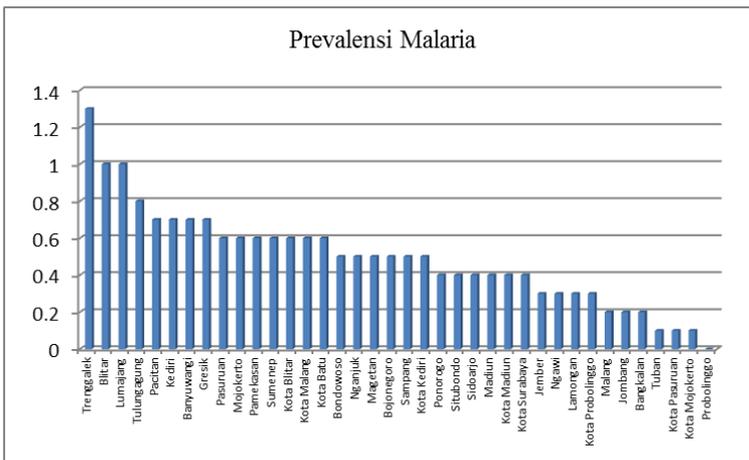


Gambar 4.5 Grafik Prevalensi TBC di Jawa Timur

Nilai rata-rata prevalensi TBC di Jawa Timur adalah sebesar 0,2263%. Nilai tersebut berarti bahwa terdapat 0,2263% dari seluruh penduduk di Jawa Timur yang menderita penyakit TBC. Nilai rata-rata tersebut masih berada di bawah rata-rata nasional yaitu sebesar 0,4%. Nilai minimum prevalensi TBC di Jawa Timur adalah 0% yang artinya pada kabupaten/kota tersebut tidak terdapat penduduk yang didiagnosa menderita TBC. Namun bukan berarti pada daerah tersebut tidak terdapat penduduk yang terkena TBC, sebab terdapat penduduk di daerah tersebut yang memiliki gejala TBC namun masih belum terdiagnosa oleh tenaga kesehatan. Terdapat tiga wilayah di Jawa Timur yang memiliki prevalensi TBC sebesar 0%, yaitu Kabupaten Bondowoso, Pamekasan dan Kota Blitar. Berdasarkan data Dinas Kesehatan di Jawa Ti-

mur, kecenderungan diagnosis kasus TBC semakin meningkat seiring dengan meningkatnya kelompok umur. Selain itu penduduk laki-laki dengan pendidikan di bawah SD dan tinggal di daerah perkotaan lebih cenderung untuk dididagnosis menderita TBC.

Gambar 4.6 yang menggambarkan prevalensi malaria di Jawa Timur adalah sebagai berikut:



Gambar 4.6 Grafik Prevalensi Malaria di Jawa Timur

Rata-rata prevalensi malaria di Jawa Timur mencapai 0,4895%. Nilai rata-rata tersebut masih berada di bawah rata-rata prevalensi malaria nasional yaitu sebesar 0,6%. Berdasarkan Gambar 4.6 terlihat bahwa terdapat daerah yang tidak ditemukan penderita yang didiagnosis menderita penyakit malaria, prevalensi 0% yaitu Kabupaten Probolinggo, namun terdapat daerah yang memiliki prevalensi malaria mencapai 1,3%, yaitu Kabupaten Trenggalek. Rentang tingkat prevalensi yang cukup besar tersebut perlu mendapatkan perhatian pemerintah. Penyakit malaria merupakan masalah kesehatan masyarakat karena sering menimbulkan kejadian Luar Biasa (KLB), berdampak luas terhadap kualitas hidup dan ekonomi, serta dapat mengakibatkan kematian.

4.2 Mereduksi Variabel Penelitian

Setelah dilakukan deskripsi data menggunakan statistika deskriptif, langkah selanjutnya adalah menentukan jumlah kluster optimum. Namun sebelum menentukan jumlah kluster optimum dengan menggunakan nilai *Pseudo F-statistics* perlu dilakukan pengujian Bartlett yang bertujuan untuk mengetahui ada atau tidaknya hubungan di antara variabel penelitian.

Sebelum dilakukan pengelompokan pada kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat terlebih dahulu dilakukan pengujian Bartlett untuk mengetahui ada atau tidaknya hubungan di antara variabel penelitian yang dilakukan secara multivariat. Antara variabel penelitian yang satu dengan yang lainnya harusnya bersifat independen atau tidak memiliki hubungan.

Hipotesis yang digunakan pada Uji Bartlett adalah sebagai berikut:

$H_0: \rho = I$ artinya tidak terdapat hubungan antar variabel

$H_1: \rho \neq I$ artinya terdapat hubungan antar variabel

Dengan statistik uji dapat dilihat pada persamaan (2.3) dan daerah penolakan adalah tolak H_0 jika $\chi^2 > \chi^2_{(1/2p)(p-1); \alpha}$ atau jika $p\text{-value} < \alpha$, yang berarti bahwa antara variabel saling berkorelasi.

Hasil pengujian Bartlett diperoleh nilai χ^2 hitung sebesar 111,572 dengan derajat bebas 15. Nilai tersebut lebih besar dibandingkan dengan nilai χ^2 tabel yaitu sebesar 24,996. Pada pengujian juga diperoleh nilai $p\text{-value}$ sebesar 0,000 yang lebih kecil dibandingkan dengan $\alpha = 0,05$. Berdasarkan kedua nilai pada hasil pengujian Bartlett dapat diambil keputusan Tolak H_0 , yang berarti bahwa terdapat hubungan atau korelasi di antara variabel penelitian.

Adanya korelasi atau hubungan antar variabel dapat menyebabkan pengelompokan menjadi tidak optimum. Untuk mengatasi hal tersebut dapat dilakukan reduksi variabel dengan menggunakan analisis faktor. Variabel yang saling berkorelasi akan dikelompokkan ke dalam satu faktor, sehingga dari beberapa variabel

akan dibentuk menjadi beberapa faktor dimana masing-masing faktor tersebut bersifat independen.

Berikut ini adalah *eigen value* yang dapat digunakan untuk menentukan banyaknya faktor baru yang terbentuk:

Tabel 4.1 Nilai *Eigen Value* Analisis Faktor

Banyak Komponen	<i>Eigen Value</i>		
	Total	% Variasi	% Kumulatif
1	2,796	46,605	46,605
2	1,317	21,951	68,556
3	0,897	14,946	83,502
4	0,691	11,511	95,013
5	0,225	3,746	98,759
6	0,074	1,241	100

Jumlah faktor baru yang terbentuk dapat didasarkan pada banyaknya komponen yang memiliki nilai *eigen value* atau λ yang lebih besar dari nilai 1. Pada Tabel 4.1 terlihat bahwa variabel akan direduksi menjadi 2 faktor apabila didasarkan pada nilai lambda yang lebih besar dari 1. Namun dengan mereduksi variabel menjadi 2 faktor, kedua faktor tersebut hanya mampu menjelaskan sebesar 68,556% dari variasi total. Nilai tersebut masih relatif kecil, sehingga variabel akan direduksi menjadi 3 faktor. dimana ketiga faktor baru yang terbentuk akan mampu menjelaskan sebesar 83,502% dari keragaman total.

Nilai *loading factor* pada Tabel 4.2 menunjukkan bahwa variabel Angka Harapan Hidup (X_2), Angka Melek Huruf (X_3) dan Rata-rata Lama Sekolah (X_4) dikelompokkan menjadi 1 faktor sebab ketiga variabel tersebut memiliki nilai korelasi yang ketika pada komponen 1. Variabel prevalensi malaria (X_6) diletakkan ke dalam faktor 3. Namun variabel PDRB per Kapita (X_1) dan Prevalensi TBC (X_5) masih sulit untuk ditentukan akan diletakkan pada faktor yang mana. Oleh sebab itu dilakukan rotasi untuk mempermudah pengelompokan variabel dan penamaan faktor

yang terbentuk. Jenis rotasi yang digunakan adalah rotasi *Vari-max*. Berikut ini adalah nilai *loading factor* untuk menentukan pengelompokan variabel menjadi beberapa faktor yang baru:

Tabel 4.2 Nilai *Loading Factor* untuk Reduksi Variabel

Variabel	Jumlah Komponen		
	1	2	3
PDRB per Kapita (X_1)	0,412	0,586	0,328
Angka Harapan Hidup (X_2)	0,887	-0,216	0,056
Angka Melek Huruf (X_3)	0,948	0,061	-0,059
Rata-rata Lama Sekolah (X_4)	0,922	0,194	-0,183
Prevalensi TBC (X_5)	-0,226	0,684	0,477
Prevalensi Malaria (X_6)	0,204	-0,647	0,722

Rotasi nilai *loading factor* disajikan dalam Tabel 4.4 sebagai berikut:

Tabel 4.3 Rotasi Nilai *Loading Factor* untuk Reduksi Variabel

Variabel	Komponen		
	1	2	3
PDRB per Kapita (X_1)	0,397	0,680	-0,034
Angka Harapan Hidup (X_2)	0,856	-0,107	0,303
Angka Melek Huruf (X_3)	0,949	0,046	0,055
Rata-rata Lama Sekolah (X_4)	0,948	0,076	-0,126
Prevalensi TBC (X_5)	-0,247	0,824	-0,069
Prevalensi Malaria (X_6)	0,064	-0,073	0,986

Setelah dilakukan rotasi pada nilai *loading factor* diperoleh hasil bahwa variabel Angka Harapan Hidup (X_2), Angka Melek Huruf (X_3) dan Rata-rata Lama Sekolah (X_4) dikelompokkan menjadi 1 faktor dan nama faktor pembentuk IPM (Indeks Pembangunan Manusia). Variabel prevalensi malaria (X_6) diletakkan ke dalam faktor 3 dengan nama prevalensi malaria sebab hanya

terdapat satu variabel pada faktor tersebut. Variabel PDRB per Kapita (X_1) dan prevalensi TBC (X_5) diletakkan ke dalam faktor 2 dengan nama variabel ekonomi.

Pada analisis faktor akan diperoleh nilai *score factor* yang merupakan kombinasi antara variabel yang direduksi ke dalam beberapa faktor. Pada penelitian ini enam variabel direduksi menjadi tiga faktor, sehingga terdapat tiga *score factor* yang akan digunakan sebagai nilai pengganti enam variabel. Pengelompokan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat akan dilakukan berdasarkan tiga nilai *score factor* yang terbentuk.

4.3 Pemilihan Kluster Terbaik

Pengelompokan kabupaten/kota menggunakan metode Kohonen SOM dan *K-Means* masing-masing akan dibentuk menjadi 2 kluster dan 3 kluster. Langkah pertama adalah mengelompokan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat menggunakan metode Kohonen SOM. Dimana pada pengelompokan menggunakan metode Kohonen SOM akan digunakan tiga jenis topologi, yaitu *gridtop*, *hextop*, dan *randtop*. Berikut adalah nilai *Pseudo F-statistics* pada masing-masing kelompok dan jenis topologi metode Kohonen SOM:

Tabel 4.4 Nilai *Pseudo F-statistics* untuk Metode Kohonen SOM

Jumlah Neuron	Jenis Topologi		
	<i>Gridtop</i>	<i>Hextop</i>	<i>Randtop</i>
2 kluster	11,054	11,967	11,928
3 kluster	13,814	13,819	13,819

Untuk menentukan jumlah kelompok optimum dapat digunakan kriteria nilai *Pseudo F-statistics*. Kelompok optimum merupakan kelompok yang memiliki nilai *Pseudo F-statistics* yang semakin tinggi. Pada Tabel 4.4 terlihat bahwa ketika digunakan jumlah kelompok sebanyak 2 kluster maka diperoleh nilai *Pseudo F-statistics* tertinggi sebesar 11,967. Dimana nilai *Pseudo*

F-statistics tersebut merupakan hasil pengelompokan metode Kohonen SOM menggunakan jenis topologi *hextop*.

Hasil yang sama diperoleh ketika digunakan jumlah kelompok sebanyak 3 klaster. Nilai *Pseudo F-statistics* tertinggi dihasilkan pada metode Kohonen SOM ketika menggunakan jenis topologi *hextop* yaitu sebesar 13,819. Informasi lain yang diperoleh pada Tabel 4.4 adalah nilai *Pseudo F-statistics* ketika digunakan jumlah kelompok sebanyak 3 klaster lebih besar dibandingkan dengan ketika digunakan jumlah kelompok sebanyak 2 klaster. Oleh sebab itu pengelompokan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat akan dikelompokkan menjadi 3 klaster menggunakan metode Kohonen SOM dengan jenis topologi yang digunakan adalah *hextop*.

Selanjutnya dilakukan pengelompokan menggunakan metode *K-Means*. Jumlah kelompok yang digunakan pada metode *K-Means* sama dengan jumlah kelompok yang digunakan pada metode Kohonen SOM, yaitu sebanyak 2 klaster dan 3 klaster. Ketika digunakan jumlah kelompok sebanyak 2 klaster diperoleh nilai *Pseudo F-statistics* sebesar 9,781 sedangkan ketika digunakan jumlah kelompok sebanyak 3 klaster diperoleh nilai *Pseudo F-statistics* sebesar 2,793. Nilai *Pseudo F-statistics* ketika digunakan jumlah kelompok 2 klaster lebih besar dibandingkan dengan ketika digunakan jumlah kelompok sebanyak 3 klaster. Berdasarkan hal tersebut maka pengelompokan menggunakan metode *K-Means* akan dilakukan dengan jumlah kelompok sebanyak 2 klaster.

Setelah diperoleh jumlah kelompok optimum untuk mengelompokkan kabupaten/kota di Jawa Timur pada masing-masing metode, langkah selanjutnya adalah menentukan hasil kelompok terbaik berdasarkan kriteria nilai *icdrate*. Nilai *icdrate* (*internal cluster dispersion rate*) merupakan nilai yang menunjukkan tingkat perbedaan atau dispersi antar anggota dalam satu klaster. Apabila suatu kelompok memiliki nilai *icdrate* suatu kelompok yang semakin kecil, maka dapat dikatakan bahwa hasil pengelompokan tersebut telah baik. Hal tersebut dikarenakan nilai *icdrate* antar

anggota dalam satu kelompok yang semakin kecil menunjukkan bahwa perbedaan dalam kelompok tersebut semakin kecil atau dengan kata lain hasil pengelompokan semakin homogen.

Berikut ini adalah nilai *icdrate*, SSW (*Sum Square Within*) dan SSB (*Sum Square Between*) hasil pengelompokan menggunakan metode Kohonen SOM dan *K-Means*.

Tabel 4.5 Nilai *Icdrate*, SSW dan SSB untuk Metode Kohonen SOM dan *K-Means*

Metode	Jumlah Kelompok	<i>icdrate</i>	SSW	SSB
Kohonen SOM	3 klaster	0,962	62,023	4,261
<i>K-Means</i>	2 klaster	0,988	87,283	1,312

Pengelompokan kabupaten/kota di Jawa Timur menggunakan metode Kohonen SOM dengan membagi data menjadi 3 klaster memiliki nilai *icdrate* yang lebih kecil dibandingkan dengan mengelompokkan kabupaten/kota di Jawa Timur menjadi 2 klaster menggunakan metode *K-Means*. Berdasarkan hal tersebut dalam disimpulkan bahwa hasil klaster terbaik adalah menggunakan metode Kohonen SOM dengan jumlah kelompok sebanyak 3 klaster.

Nilai SSW (*Sum Square Within*) merupakan nilai jarak total antar anggota dalam satu klaster dengan pusat klasternya sedangkan nilai SSB (*Sum Square Between*) merupakan nilai jarak total antar pusat klaster. Klaster terbaik adalah klaster dengan kondisi antar anggota klaster dalam satu klaster yang semakin homogen sedangkan antar klaster memiliki karakteristik yang semakin heterogen. Dengan kata lain, klaster terbaik adalah klaster yang memiliki nilai SSW yang semakin kecil dan nilai SSB yang semakin besar.

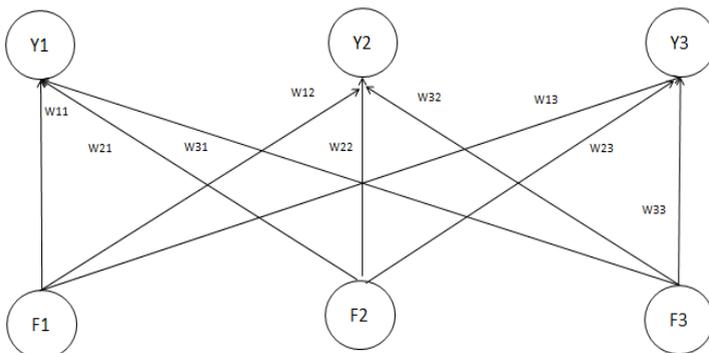
Berdasarkan Tabel 4.5 diketahui bahwa mengelompokkan kabuapten/kota di Jawa Timur menjadi 3 klaster menggunakan Kohonen SOM memiliki nilai SSW yang lebih kecil dibandingkan dengan metode *K-Means*. Selain itu nilai SSB metode Kohonen SOM lebih besar dibandingkan dengan nilai SSB metode *K-*

Means. Kedua hal tersebut juga menunjukkan bahwa hasil kluster terbaik adalah menggunakan metode Kohonen SOM dengan jumlah kelompok sebanyak 3 kluster.

4.4 Pengelompokan Menggunakan Metode Terbaik

Berdasarkan kriteria nilai *Pseudo F-statistics* dan nilai *icd-rate* diketahui bahwa hasil kluster terbaik merupakan pengelompokan menggunakan metode Kohonen SOM dengan topologi *hextop* dan jumlah neuron sebanyak 3. Jumlah neuron pada jaringan Kohonen SOM menunjukkan banyaknya kelompok yang terbentuk. Dengan menggunakan jumlah neuron sebanyak 3 maka jaringan arsitektur Kohonen SOM yang sesuai untuk mengelompokkan kabupaten/kota di Jawa Timur disajikan pada Gambar 4.7.

Terdapat 3 vektor *input*, dimana 3 faktor tersebut merupakan hasil reduksi variabel yang mewakili 6 variabel awal. Masing-masing vektor *input* i yang menuju neuron ke j dihubungkan dengan bobot. Bobot tersebut merupakan nilai *centroid* atau pusat kluster.



Gambar 4.7 Arsitektur Jaringan Kohonen SOM Menggunakan 3 Neuron

Hasil bobot akhir yang telah optimum untuk mengelompokkan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat dapat disajikan dalam bentuk matriks bobot berukuran 3 x 3 sebagai berikut:

$$w_{ij} = \begin{pmatrix} -1,1036 & 0,0617 & -0,1767 \\ 0,6565 & -0,4216 & -0,4160 \\ 0,2248 & 0,5007 & 1,3204 \end{pmatrix} -$$

Hasil pengelompokan yang terbentuk pada kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat secara visual dapat dilihat pada gambar sebagai berikut:



Gambar 4.8 Hasil Pengelompokan Kluster Terbaik

Dengan menggunakan jumlah kelompok sebanyak 3 kluster hasil pengelompokan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat adalah sebagai berikut kluster 1 terdiri dari 12 kabupaten yang meliputi Kabupaten Jember, Bondowoso, Situbondo, Probolinggo, Pasuruan, Bojonegoro, Tuban, Lamongan, Bangkalan, Sampang, Pamekasan, Sumenep. Kluster 2 terdiri dari 17 kabupaten/kota yang meliputi Ponorogo, Malang, Sidoarjo, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota

Batu. Serta klaster 3 yang terdiri dari 5 kabupaten/kota yaitu Pacitan, Trenggalek, Tulungagung, Blitar, Kediri, Lumajang, Banyuwangi, GrasiK, Kota Surabaya.

Berikut ini adalah nilai rata-rata untuk setiap variabel pada masing-masing kelompok yang terbentuk:

Tabel 4.6 Nilai Rata-rata Hasil Pengelompokan Klaster Terbaik

Variabel	Kelompok 1	Kelompok 2	Kelompok 3	Rata-rata Populasi
PDRB	6315,70	8356,6	19538,5	10360
Angka Harapan Hidup	64,88	70,5	70,7	68,763
Angka Melek Huruf	82,40	94,1	92,6	90,03
Lama Sekolah	6,21	8,7	7,8	7,695
Prevalensi TBC	0,28	0,17	0,27	0,2263
Prevalensi Malaria	0,38	0,39	0,81	0,4895

Nilai rata-rata populasi merupakan nilai rata-rata setiap variabel secara keseluruhan. Untuk mengetahui karakteristik dari klaster yang terbentuk dapat digunakan nilai rata-rata dari masing-masing variabel pada setiap klaster yang kemudian dibandingkan dengan nilai rata-rata setiap variabel secara keseluruhan. Klaster 1 merupakan kelompok yang memiliki nilai rata-rata PDRB per kapita lebih rendah dibandingkan dengan nilai rata-rata PDRB populasi. Selain itu klaster 1 juga memiliki nilai rata-rata Angka Harapan Hidup, Angka Melek Huruf dan lama sekolah yang berada di bawah rata-rata populasi. Berdasarkan hal tersebut dapat disimpulkan bahwa klaster 1 merupakan kelompok yang memiliki kondisi ekonomi dan pendidikan yang kurang baik. Meskipun demikian klaster 1 merupakan kelompok yang memiliki kondisi epidemiologi yang cukup baik. Hal tersebut dapat diketahui berdasarkan nilai rata-rata prevalensi malaria yang berada di bawah nilai rata-rata populasi. Namun klaster 1 merupakan ke-

lompok yang memiliki prevalensi TBC yang berada di atas nilai rata-rata prevalensi TBC populasi.

Klaster 2 merupakan klaster 2 merupakan kelompok yang memiliki nilai rata-rata PDRB per kapita yang berada di bawah nilai rata-rata PDRB per kapita populasi namun memiliki nilai rata-rata Angka Harapan Hidup yang berada di atas nilai rata-rata Angka Harapan Hidup Populasi. Berdasarkan hal tersebut dapat dikatakan bahwa klaster 2 merupakan kelompok dengan kondisi ekonomi yang cukup baik. Klaster 2 memiliki nilai rata-rata Angka Melek Huruf dan lama sekolah yang berada di atas nilai rata-rata populasi. Oleh sebab itu dapat disimpulkan bahwa klaster 2 merupakan kelompok yang memiliki kondisi pendidikan yang baik. Selain itu pada variabel epidemiologi, diketahui bahwa klaster 2 memiliki nilai rata-rata prevalensi TBC dan malaria yang berada di bawah nilai rata-rata populasi. Berdasarkan hal tersebut dikatakan bahwa klaster 2 merupakan kelompok yang memiliki kondisi epidemiologi yang baik.

Klaster 3 memiliki nilai rata-rata PDRB per kapita dan Angka Harapan Hidup yang lebih tinggi dibandingkan dengan nilai rata-rata populasi. Oleh sebab itu klaster 3 merupakan kelompok dengan kondisi ekonomi yang baik. Selain itu klaster 3 juga merupakan kelompok yang memiliki kondisi pendidikan yang baik. Hal tersebut dapat terlihat dari nilai rata-rata Angka Melek Huruf dan lama sekolah pada klaster 3 lebih tinggi dibandingkan dengan nilai rata-rata populasi. Namun meskipun memiliki kondisi ekonomi dan pendidikan yang baik, klaster 3 merupakan kelompok yang memiliki kondisi epidemiologi yang kurang baik. Kondisi tersebut tercermin pada nilai rata-rata prevalensi TBC dan malaria pada klaster 3 lebih tinggi dibandingkan dengan nilai rata-rata populasi.

Berdasarkan uraian sebelumnya dapat disimpulkan bahwa klaster 1 merupakan klaster yang memiliki kondisi ekonomi dan pendidikan yang kurang baik, namun memiliki kondisi epidemiologi yang cukup baik. Klaster 2 memiliki kondisi ekonomi yang cukup baik namun memiliki kondisi pendidikan dan epidemiologi

yang baik. Klaster 3 merupakan kelompok yang memiliki kondisi ekonomi dan pendidikan yang baik namun memiliki kondisi epidemiologi yang kurang baik.

Untuk meningkatkan kondisi kesehatan masyarakat di Jawa Timur, monitoring dan evaluasi dapat dilakukan oleh Pemerintah Jawa Timur berdasarkan klaster yang terbentuk. Pada klaster 1 yang dapat disebut sebagai klaster yang memiliki kondisi kesehatan masyarakat yang paling rendah, pemerintah perlu memperbaiki kondisi ekonomi, pendidikan dan epidemiologi. Sedangkan pada klaster 3 yang dapat disebut sebagai klaster dengan kondisi ekonomi yang sedang, pemerintah perlu memperbaiki kondisi epidemiologi yang bertujuan untuk mengurangi angka prevalensi penyakit TBC dan malaria. Selain itu pemerintah juga perlu mempertahankan kondisi ekonomi dan pendidikan yang berada pada klaster 3 sebab klaster 3 telah memiliki kondisi ekonomi dan pendidikan yang baik. Klaster 2 dapat disebut sebagai kelompok yang memiliki kondisi kesehatan masyarakat yang tinggi. Dimana hal tersebut tercermin pada kondisi ekonomi, pendidikan dan epidemiologi pada klaster 2 yang telah baik. Meskipun demikian, pemerintah perlu menjaga kondisi kesehatan masyarakat pada klaster 2 dan juga meningkatkan nilai PDRB per kapita pada klaster 2.

Pengelompokan pada kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat bertujuan untuk mendapatkan kelompok-kelompok kabupaten/kota yang memiliki karakteristik antar anggota dalam satu klaster yang sama namun memiliki perbedaan antar kelompok yang terbentuk. Hasil kelompok kabupaten/kota di Jawa Timur menggunakan metode Kohonen SOM diharapkan memiliki perbedaan karakteristik antar kelompok yang terbentuk. Untuk membuktikan hal tersebut dapat dilakukan pengujian MANOVA (*Multivariate Analysis of Variance*) untuk mengetahui ada atau tidaknya perbedaan pada setiap klaster yang terbentuk. Pengujian MANOVA yang akan dilakukan merupakan pengujian *One-way* MANOVA dimana hanya terdapat satu faktor

atau perlakuan dan tidak terdapat interaksi antar perlakuan yang diduga mempengaruhi variabel respon.

Pada pengujian *One-way* MANOVA, faktor atau perlakuan yang diduga memberikan perbedaan pada variabel respon adalah kelompok yang terbentuk. Dimana satu faktor tersebut terdiri dari tiga kategori, yaitu kelompok 1, kelompok 2 dan kelompok 3. Variabel respon pada pengujian *One-way* MANOVA adalah PDRB per kapita, Angka Harapan Hidup, Angka Melek Huruf, lama sekolah, prevalensi TBC dan prevalensi malaria.

Hasil pengujian *One-way* MANOVA disajikan pada Tabel 4.7 sebagai berikut:

Tabel 4.7 Hasil Pengujian *One-way* MANOVA pada Klaster

Variabel	<i>P-value</i>
PDRB	0,194
Angka Harapan Hidup	0,000
Angka Melek Huruf	0,000
Lama Sekolah	0,000
Prevalensi TBC	0,092
Prevalensi Malaria	0,000

Hipotesis yang digunakan pada pengujian *One-way* MANOVA adalah sebagai berikut:

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \dots = \tau_g = 0$$

$$H_1 : \text{minimal ada 1 } \tau_g \neq 0, \text{ dengan } l = 1, 2, \dots, g$$

Berdasarkan Tabel 4.7 terlihat bahwa pada hasil pengujian *One-way* MANOVA diperoleh nilai *p-value* untuk variabel Angka Harapan Hidup, Angka Melek Huruf, lama sekolah dan prevalensi malaria masing-masing sebesar 0,000. Nilai tersebut lebih kecil dibandingkan dengan nilai $\alpha = 0,05$ sehingga Tolak H_0 . Keputusan Tolak H_0 berarti bahwa variabel Angka Harapan Hidup, Angka Melek Huruf, lama sekolah dan prevalensi malaria pada masing-masing klaster 1, klaster 2 dan klaster 3 yang terbentuk memiliki perbedaan. Sedangkan variabel PDRB per kapita memiliki nilai *p-value* sebesar 0,194 dan prevalensi TBC memiliki nilai *p-value*

sebesar 0,092. Kedua nilai tersebut lebih besar dibandingkan dengan nilai $\alpha = 0,05$ sehingga Gagal Tolak H_0 , yang artinya variabel PDRB per kapita dan prevalensi TBC pada masing-masing klaster yang terbentuk adalah tidak memiliki perbedaan.

Pada hasil pengujian *One-way* MANOVA diperoleh kesimpulan bahwa hasil pengelompokan memiliki perbedaan nilai pada masing-masing variabel respon yang membentuknya. Meskipun terdapat dua variabel yang tidak memiliki perbedaan pada setiap klaster yang terbentuk, yaitu PDRB per kapita dan prevalensi TBC. Namun ketika digunakan nilai $\alpha = 0,1$ maka variabel prevalensi TBC sebab nilai *p-value* pada variabel TBC sebesar 0,092 lebih kecil dibandingkan dengan nilai α . Sehingga dapat disimpulkan bahwa variabel prevalensi TBC memiliki perbedaan pada masing-masing klaster yang terbentuk.

Berdasarkan uraian tersebut maka lebih disarankan untuk menggunakan nilai $\alpha = 0,1$ pada saat melakukan pengujian *One-way* MANOVA. Sebab dengan menggunakan nilai nilai $\alpha = 0,1$ maka hanya terdapat satu variabel yang tidak memberikan beda pada setiap klaster yang terbentuk, yaitu variabel PDRB per kapita.

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Setelah dilakukan analisis dan pembahasan mengenai pengelompokan kabupaten/kota di Jawa Timur menggunakan metode Kohonen SOM dan *K-Means* dapat diambil kesimpulan sebagai berikut:

1. Hasil analisis statistika deskriptif adalah rentang nilai maksimum dan minimum pada variabel PDRB per kapita di Jawa Timur yang cukup besar mengindikasikan bahwa pertumbuhan ekonomi di Jawa Timur masih belum merata pada masing-masing kabupaten/kota. Nilai rata-rata Angka Harapan Hidup di Jawa Timur pada tahun 2012 adalah sebesar 68,76 tahun. Nilai rata-rata tersebut masih berada di bawah rata-rata Angka Harapan Hidup nasional yaitu sebesar 69,87 tahun. Rata-rata Angka Melek Huruf sebesar 90,03 yang masih berada di bawah rata-rata Angka Melek Huruf nasional di tahun 2012 yaitu sebesar 93,25. Nilai rata-rata lama sekolah terbesar di Jawa Timur dimiliki oleh Kota Malang yaitu sebesar 10,87 tahun sedangkan nilai rata-rata terendah dimiliki oleh Kabupaten Sampang yaitu sebesar 4,22 tahun. Nilai rata-rata prevalensi TBC di Jawa Timur adalah sebesar 0,2263%. Nilai tersebut berarti bahwa terdapat 0,2263% dari seluruh penduduk di Jawa Timur yang menderita penyakit TBC. Rata-rata prevalensi malaria di Jawa Timur mencapai 0,4895%. Nilai rata-rata tersebut masih berada di bawah rata-rata prevalensi malaria nasional yaitu sebesar 0,6%.
2. Jumlah faktor baru yang terbentuk pada hasil reduksi variabel adalah sebanyak 3 faktor. Faktor 1 yaitu faktor ekonomi terdiri dari variabel PDRB per kapita dan prevalensi TBC, faktor 2 yaitu faktor pembentuk IPM yang terdiri dari variabel Angka Harapan Hidup, Angka Melek Huruf dan lama sekolah serta faktor 3 yaitu faktor prevalensi malaria. Pada metode Kohonen

SOM diperoleh nilai *Pseudo F-statistics* terbesar yaitu 13,819 ketika digunakan jumlah kelompok sebanyak 3 klaster dengan jenis topologi *hextop*. Sedangkan pada metode *K-Means* nilai *Pseudo F-statistics* terbesar yaitu 9,781 ketika digunakan kelompok sebanyak 2 klaster. Perbandingan hasil klaster terbaik berdasarkan nilai *icdrate* diperoleh kesimpulan bahwa nilai *icdrate* metode Kohonen SOM yaitu sebesar 0,962 lebih kecil dibandingkan dengan nilai *icdrate* hasil pengelompokan metode *K-Means* yaitu sebesar 0,988.

3. Monitoring dan evaluasi dapat dilakukan oleh Pemerintah Jawa Timur berdasarkan klaster yang terbentuk. Pada klaster 1 yang dapat disebut sebagai klaster yang memiliki kondisi kesehatan masyarakat yang paling rendah, pemerintah perlu memperbaiki kondisi ekonomi, pendidikan dan epidemiologi. Sedangkan pada klaster 3 yang dapat disebut sebagai klaster dengan kondisi ekonomi yang sedang, pemerintah perlu memperbaiki kondisi epidemiologi yang bertujuan untuk mengurangi angka prevalensi penyakit TBC dan malaria. Selain itu pemerintah juga perlu mempertahankan kondisi ekonomi dan pendidikan yang berada pada klaster 3 sebab klaster 3 telah memiliki kondisi ekonomi dan pendidikan yang baik. Klaster 2 dapat disebut sebagai kelompok yang memiliki kondisi kesehatan masyarakat yang tinggi. Dimana hal tersebut tercermin pada kondisi ekonomi, pendidikan dan epidemiologi pada klaster 2 yang telah baik. Meskipun demikian, pemerintah perlu menjaga kondisi kesehatan masyarakat pada klaster 2 dan juga meningkatkan nilai PDRB per kapita pada klaster 2.

5.2 Saran

Saran yang dapat diberikan kepada Pemerintah Provinsi Jawa Timur dan Dinas Kesehatan Provinsi Jawa Timur adalah hasil penelitian ini dapat digunakan sebagai informasi untuk meningkatkan kondisi kesehatan masyarakat di Jawa Timur dalam rangka mencapai Program Kota Sehat Nasional Pemerintah perlu memperhatikan variabel ekonomi dan pendidikan, yaitu PDRB

per kapita, Angka Harapan Hidup, Angka Melek Huruf dan Rata-rata lama sekolah. Kondisi ekonomi dan pendidikan di Jawa Timur yang semakin baik dapat menciptakan kondisi kesehatan masyarakat yang lebih baik.

Saran untuk penelitian selanjutnya adalah menggunakan variabel penelitian yang mempresentasikan kondisi kesehatan masyarakat di Jawa Timur secara menyeluruh sehingga hasil pengelompokan yang dihasilkan juga optimal.

(Halaman ini sengaja dikosongkan)

LAMPIRAN

LAMPIRAN A DATA PENELITIAN

Kota	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
Pacitan	1695,28	71,69	91,63	6,96	0,3	0,7
Ponorogo	3727,28	70,4	88,99	7,18	0,1	0,4
Trenggalek	3394,02	72,13	92,88	7,31	0,3	1,3
Tulungagung	9065,39	71,95	94,57	7,95	0,1	0,8
Blitar	6371,39	71,3	92,05	7,4	0,1	1
Kediri	8894,01	70,15	92,87	7,72	0,4	0,7
Malang	17441,76	69,5	90,73	7,08	0,2	0,2
Lumajang	7285,78	67,75	86,58	6,43	0,2	1
Jember	13371,42	63,21	83,65	6,79	0,4	0,3
Banyuwangi	12657,00	68,38	88,08	7,25	0,2	0,7
Bondowoso	3672,61	63,85	80,72	5,94	0	0,5
Situbondo	4006,10	63,52	78,31	6,22	0,4	0,4
Probolinggo	7894,21	61,7	80,48	5,92	0,3	0
Pasuruan	7943,83	64,61	91,17	6,83	0,4	0,6
Sidoarjo	32215,99	71,03	97,79	9,92	0,1	0,4
Mojokerto	9623,85	70,64	94,16	7,94	0,1	0,6
Jombang	7172,85	70,28	93,87	8,04	0,2	0,2
Nganjuk	6047,18	69,33	91,11	7,61	0,1	0,5
Madiun	3381,79	69,25	89,61	7,44	0,2	0,4
Magetan	3581,55	71,66	91,08	7,85	0,2	0,5
Ngawi	3420,66	70,57	85,58	7,02	0,2	0,3
Bojonegoro	7933,29	67,42	84,85	6,72	0,2	0,5
Tuban	9057,89	68,21	85,86	6,53	0,2	0,1
Lamongan	6836,69	68,55	88,76	7,59	0,4	0,3
Gresik	19424,16	71,47	96,17	8,98	0,4	0,7
Bangkalan	3834,45	63,65	82,9	5,74	0,5	0,2
Sampang	3275,58	63,98	69,12	4,22	0,1	0,5
Pamekasan	2401,24	64,79	84,21	6,32	0	0,6
Sumenep	5561,06	65,07	78,71	5,71	0,4	0,6
Kota Kediri	27590,86	70,86	97,6	10,24	0,1	0,5
Kota Blitar	1155,93	72,8	97,31	9,77	0	0,6

LAMPIRAN A
DATA PENELITIAN (LANJUTAN)

Kota	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
Kota Malang	17380,76	71,02	98,34	10,87	0,1	0,6
Kota Probolinggo	2367,47	70,86	92,55	8,67	0,5	0,3
Kota Pasuruan	1310,47	66,46	97,07	9,05	0,3	0,1
Kota Mojokerto	1485,04	72	97,18	10,11	0,1	0,1
Kota Madiun	2474,02	71,42	97,84	10,46	0,3	0,4
Kota Surabaya	107059,48	71,53	98,35	10,1	0,4	0,4
Kota Batu	1684,12	70	98,32	8,54	0,1	0,6

Keterangan:

X₁ : Pendapatan Domestik Regional Bruto (PDRB) per kapita menurut harga konstan

X₂ : Angka Harapan Hidup

X₃ : Angka Melek Huruf

X₄ : Rata-rata lama sekolah

X₅ : Prevalensi TBC

X₆ : Prevalensi malaria

LAMPIRAN B

HASIL ANALISIS FAKTOR

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,651
Bartlett's Test of Sphericity	Approx. Chi-Square	111,572
	df	15
	Sig.	,000

Communalities		
	Initial	Extraction
X1	1,000	,621
X2	1,000	,836
X3	1,000	,905
X4	1,000	,921
X5	1,000	,745
X6	1,000	,982
Extraction Method: Principal Component Analysis.		

Total Variance Explained			
Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2,796	46,605	46,605
2	1,317	21,951	68,556
3	,897	14,946	83,502
4	,691	11,511	95,013
5	,225	3,746	98,759
6	,074	1,241	100,000
Extraction Method: Principal Component Analysis.			

Component Matrix^a			
	Component		
	1	2	3
X1	,412	,586	,328
X2	,887	-,216	,056
X3	,948	,061	-,059
X4	,922	,194	-,183
X5	-,226	,684	,477
X6	,204	-,647	,722
Extraction Method: Principal Component Analysis			
a. 3 components extracted			

Rotated Component Matrix^a			
	Component		
	1	2	3
X1	,397	,680	-,034
X2	,856	-,107	,303
X3	,949	,046	,055
X4	,948	,076	-,126
X5	-,247	,824	-,069
X6	,064	-,073	,986
Extraction Method: Principal Component Analysis.			
Rotation Method: Varimax with Kaiser Normalization.			
a. Rotation converged in 4 iterations.			

LAMPIRAN C

NILAI SCORE FACTOR

Kabupaten/Kota	Faktor 1	Faktor 2	Faktor 3
Pacitan	0,006441	0,090744	1,031966
Ponorogo	0,064328	-0,94414	-0,2384
Trenggalek	0,052824	0,354585	3,106245
Tulungagung	0,60783	-0,65374	1,141416
Blitar	0,221163	-0,65907	1,875533
Kediri	0,067929	0,885113	0,915393
Malang	0,104932	-0,02548	-0,88832
Lumajang	-0,69446	-0,03514	1,874474
Jember	-1,13428	1,053331	-0,73702
Banyuwangi	-0,25373	0,022756	0,765747
Bondowoso	-1,19923	-1,26776	-0,16189
Situbondo	-1,59878	0,762907	-0,29362
Probolinggo	-1,51717	0,283803	-1,82886
Pasuruan	-0,72802	0,955778	0,345584
Sidoarjo	1,408912	0,00206	-0,527
Mojokerto	0,516424	-0,67401	0,36697
Jombang	0,490682	-0,37755	-1,00317
Nganjuk	0,162651	-0,79946	-0,01511
Madiun	-0,03875	-0,41173	-0,28101
Magetan	0,333905	-0,42814	0,157393
Ngawi	-0,19119	-0,48739	-0,47628
Bojonegoro	-0,61337	-0,18663	0,080345
Tuban	-0,42383	-0,31344	-1,24399
Lamongan	-0,23297	0,708587	-0,54944
Gresik	0,727405	1,207634	0,843405
Bangkalan	-1,50358	1,198402	-0,87585
Sampang	-2,26398	-0,79307	0,185304
Pamekasan	-0,87992	-1,29219	0,17744
Sumenep	-1,59286	0,844966	0,568377
Kota Kediri	1,401109	-0,10946	-0,24026
Kota Blitar	1,334228	-1,51156	0,16253
Kota Malang	1,505204	-0,41178	0,009202
Kota Probolinggo	0,312483	1,022143	-0,50891
Kota Pasuruan	0,430394	0,011634	-1,68147
Kota Mojokerto	1,376303	-1,14682	-1,60912
Kota Madiun	1,198031	0,035735	-0,51253
Kota Surabaya	1,790948	4,000684	-0,14752
Kota Batu	0,751958	-0,91229	0,212445

LAMPIRAN D**NILAI PSEUDO F, ICD-RATE, SSW, SSB**

Topologi gridtop 2 klaster

Percobaan ke	Topologi	F	<i>Icd-rate</i>	SSW	SSB
1	(1,2)	9,7681	0,9877	87,31	1,2537
	(2,1)	9,7124	0,9882	87,416	1,3047
2	(1,2)	11,0535	0,9849	84,9573	1,6743
	(2,1)	11,0535	0,9849	84,9573	1,6743
3	(1,2)	9,7728	0,9886	87,396	1,270
	(2,1)	9,7191	0,9905	87,403	1,0517
4	(1,2)	11,0535	0,9849	84,9573	1,6743
	(2,1)	10,948	0,9877	85,115	1,3699
5	(1,2)	1,6893	0,9976	106,02	0,2663
	(2,1)	110,97	0,9871	85,76	1,4341
6	(1,2)	9,7681	0,9877	87,31	1,2537
	(2,1)	10,939	0,9877	85,132	1,3614
7	(1,2)	11,0535	0,9849	84,9573	1,6743
	(2,1)	11,0535	0,9849	84,9573	1,6743
8	(1,2)	10,975	0,9876	85,067	1,3724
	(2,1)	9,6571	0,989	87,522	1,2163
9	(1,2)	10,975	0,9876	85,067	1,3724
	(2,1)	10,8627	0,9874	85,2703	1,3976
10	(1,2)	9,7472	0,9885	87,25	1,2726
	(2,1)	10,8799	0,9887	85,2309	1,2528

Topologi gridtop 3 klaster

Percobaan ke	Topologi	F	<i>Icd-rate</i>	SSW	SSB
1	(1,3)	13,8138	0,9616	62,0233	4,2607
	(3,1)	12,9514	0,9613	63,7899	4,2851
2	(1,3)	13,8138	0,9616	62,0233	4,2607
	(3,1)	12,9514	0,9613	63,7899	4,2851
3	(1,3)	13,8138	0,9616	62,0233	4,2607
	(3,1)	13,767	0,9609	61,126	4,3382
4	(1,3)	13,8138	0,9616	62,0233	4,2607
	(3,1)	13,449	0,9611	62,275	4,3182
5	(1,3)	13,8138	0,9616	62,0233	4,2607
	(3,1)	13,62	0,9624	62,42	4,1718
6	(1,3)	13,8138	0,9616	62,0233	4,2607
	(3,1)	12,951	0,9614	63,79	4,2851
7	(1,3)	13,8138	0,9616	62,0233	4,2607
	(3,1)	12,951	0,9614	63,79	4,2851
8	(1,3)	13,787	0,9609	61,126	4,3382
	(3,1)	12,951	0,9614	63,79	4,2851
9	(1,3)	13,8138	0,9616	62,0233	4,2607
	(3,1)	12,231	0,9655	65,336	3,8249
10	(1,3)	13,767	0,9609	61,126	4,3382
	(3,1)	10,274	0,9607	69,94	4,3596

Topologi hextop 2klaster

Percobaan ke	Topologi	F	Icd-rate	SSW	SSB
1	(1,2)	11,928	0,9868	83,375	1,462
	(2,1)	11,967	0,9866	83,307	1,4901
2	(1,2)	11,928	0,9868	83,375	1,462
	(2,1)	10,75	0,9827	85,476	1,9223
3	(1,2)	9,787	0,9866	87,274	1,4901
	(2,1)	11,967	0,9866	83,307	1,4901
4	(1,2)	9,787	0,9866	87,274	1,4901
	(2,1)	9,9756	0,9877	87,533	1,3611
5	(1,2)	11,928	0,9868	83,375	1,462
	(2,1)	9,7524	0,9882	87,34	1,3089
6	(1,2)	9,787	0,9866	87,274	1,4901
	(2,1)	9,7681	0,9887	87,31	1,2537
7	(1,2)	9,5957	0,9872	87,6397	1,4124
	(2,1)	9,5597	0,989	87,709	1,2558
8	(1,2)	11,831	0,9867	83,543	1,4776
	(2,1)	10,291	0,9894	86,324	1,1812
9	(1,2)	11,928	0,9868	83,375	1,462
	(2,1)	9,6905	0,9884	87,4578	1,2788
10	(1,2)	10,939	0,9877	85,132	1,3614
	(2,1)	10,2907	0,9893	86,3239	1,1812

Topologi hectop 3 klaster

Percobaan ke	Topologi	F	<i>Icd-rate</i>	SSW	SSB
1	(1,3)	12,7062	0,9626	64,3078	4,1513
	(3,1)	13,819	0,9616	62,023	4,2607
2	(1,3)	11,3471	0,9701	67,3375	3,3180
	(3,1)	12,951	0,9614	63,79	4,2851
3	(1,3)	13,819	0,9616	62,023	4,2607
	(3,1)	13,819	0,9616	62,023	4,2607
4	(1,3)	13,819	0,9616	62,023	4,2607
	(3,1)	13,819	0,9616	62,023	4,2607
5	(1,3)	13,819	0,9616	62,023	4,2607
	(3,1)	12,915	0,9641	63,866	3,9898
6	(1,3)	13,819	0,9616	62,023	4,2607
	(3,1)	12,915	0,9641	63,866	3,9898
7	(1,3)	13,819	0,9616	62,023	4,2607
	(3,1)	12,951	0,9614	63,79	4,2851
8	(1,3)	13,819	0,9616	62,023	4,2607
	(3,1)	12,951	0,9614	63,79	4,2851
9	(1,3)	12,89	0,9571	63,92	4,7661
	(3,1)	13,819	0,9616	62,023	4,2607
10	(1,3)	13,819	0,9616	62,023	4,2607
	(3,1)	13,819	0,9616	62,023	4,2607

Topologi randtop 2 klaster

Percobaan ke	Topologi	F	Icd-rate	SSW	SSB
1	(1,2)	11,928	0,9869	83,375	1,4260
	(2,1)	10,372	0,9868	85,805	1,4661
2	(1,2)	9,7870	0,9865	87,2736	1,4901
	(2,1)	11,967	0,9866	83,307	1,4901
3	(1,2)	11,928	0,9869	83,375	1,4260
	(2,1)	11,928	0,9869	83,375	1,4260
4	(1,2)	9,7870	0,9865	87,2736	1,4901
	(2,1)	10,975	0,9876	85,067	1,3724
5	(1,2)	9,7870	0,9865	87,2736	1,4901
	(2,1)	10,975	0,9876	85,067	1,3724
6	(1,2)	11,928	0,9869	83,375	1,4260
	(2,1)	11,967	0,9866	83,307	1,4901
7	(1,2)	11,928	0,9869	83,375	1,4260
	(2,1)	10,975	0,9876	85,067	1,3724
8	(1,2)	11,928	0,9869	83,375	1,4260
	(2,1)	11,928	0,9869	83,375	1,4260
9	(1,2)	9,7870	0,9865	87,2736	1,4901
	(2,1)	11,8783	0,9881	83,4615	1,3189
10	(1,2)	9,7870	0,9865	87,2736	1,4901
	(2,1)	11,8783	0,9881	83,4615	1,3189

Topologi randtop 3 klaster

Percobaan ke	Topologi	F	<i>Ic-drate</i>	SSW	SSB
1	(1,3)	13,7672	0,9609	62,1257	4,3382
	(3,1)	13,7672	0,9609	62,1257	4,3382
2	(1,3)	12,9311	0,9642	63,827	3,9641
	(3,1)	12,71	0,9638	64,3	4,0195
3	(1,3)	12,951	0,9614	63,79	4,2851
	(3,1)	13,7672	0,9609	62,1257	4,3382
4	(1,3)	12,931	0,9643	63,833	3,9641
	(3,1)	11,119	0,9682	67,875	3,5266
5	(1,3)	12,951	0,9614	63,79	4,2851
	(3,1)	13,819	0,9616	62,023	4,2607
6	(1,3)	12,951	0,9614	63,79	4,2851
	(3,1)	13,819	0,9616	62,023	4,2607
7	(1,3)	12,951	0,9614	63,79	3,9641
	(3,1)	12,71	0,9638	64,3	4,0195
8	(1,3)	13,819	0,9616	62,023	4,2607
	(3,1)	13,819	0,9616	62,023	4,2607
9	(1,3)	13,819	0,9616	62,023	4,2607
	(3,1)	13,819	0,9616	62,023	4,2607
10	(1,3)	13,688	0,9625	62,324	4,1606
	(3,1)	12,892	0,964	63,916	3,9919

LAMPIRAN E
NILAI K-MEANS

Nilai	Jumlah Klaster	
	2	3
SSW	87,283	95,722
SSB	1,312	4,260
Pseudo F	9,781	2,790
Icd-rate	0,9881	0,9616

LAMPIRAN F

PROGRAM KOHONEN SOM DI MATLAB

```
P=xlsread('D:/datafaktortrans.xlsx');  
net=newsom(P, [D1,D2], 'randtop', 'dist');  
net=train(net,P);  
b=sim(net,P);  
ac=vec2ind(b)
```

Keterangan:

P : matriks yang berisi variabel penelitian

D1, D2 : jumlah neuron pada jaringan Kohonen SOM

'randtop' : jenis topologi yang digunakan pada jaringan Kohonen SOM

'dist' : ukuran jarak yang digunakan, yaitu jarak *Euclidean*

LAMPIRAN G

OUTPUT UJI ONE-WAY MANOVA

Box's Test of Equality of Covariance Matrices^a

Box's M	112.577
F	1.920
df1	42
df2	2.215E3
Sig.	.000

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + Klaster

Multivariate Tests^c

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	1.000	1.110E4 ^a	6.000	30.000	.000	1.000
	Wilks' Lambda	.000	1.110E4 ^a	6.000	30.000	.000	1.000
	Hotelling's Trace	2.220E3	1.110E4 ^a	6.000	30.000	.000	1.000
	Roy's Largest Root	2.220E3	1.110E4 ^a	6.000	30.000	.000	1.000
Klaster	Pillai's Trace	1.411	12.389	12.000	62.000	.000	.706
	Wilks' Lambda	.084	12.221 ^a	12.000	60.000	.000	.710
	Hotelling's Trace	4.982	12.041	12.000	58.000	.000	.714
	Roy's Largest Root	3.062	15.822 ^b	6.000	31.000	.000	.754

a. Exact statistic

b. The statistic is an upper bound on F that yields a lower bound on the significance level.

c. Design: Intercept + Klaster

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	PDRB	1.023E9 ^a	2	5.114E8	1.719	.194	.089
	AHH	264.737 ^b	2	132.369	45.872	.000	.724
	AMH	1034.820 ^c	2	517.410	25.280	.000	.591
	Lama_sekolah	43.462 ^d	2	21.731	16.978	.000	.492
	Prev_TBC	.096 ^e	2	.048	2.551	.092	.127
	Prev_malaria	1.221 ^f	2	.610	14.684	.000	.456
Intercept	PDRB	4.621E9	1	4.621E9	15.536	.000	.307
	AHH	167652.756	1	167652.756	5.810E4	.000	.999
	AMH	285787.896	1	285787.896	1.396E4	.000	.997
	Lama_sekolah	2033.355	1	2033.355	1.589E3	.000	.978
	Prev_TBC	2.003	1	2.003	106.578	.000	.753
	Prev_malaria	9.964	1	9.964	239.686	.000	.873
Klaster	PDRB	1.023E9	2	5.114E8	1.719	.194	.089
	AHH	264.737	2	132.369	45.872	.000	.724
	AMH	1034.820	2	517.410	25.280	.000	.591
	Lama_sekolah	43.462	2	21.731	16.978	.000	.492
	Prev_TBC	.096	2	.048	2.551	.092	.127
	Prev_malaria	1.221	2	.610	14.684	.000	.456
Error	PDRB	1.041E10	35	2.974E8			
	AHH	100.996	35	2.886			
	AMH	716.356	35	20.467			
	Lama_sekolah	44.800	35	1.280			
	Prev_TBC	.658	35	.019			
	Prev_malaria	1.455	35	.042			
Total	PDRB	1.551E10	38				
	AHH	180042.489	38				
	AMH	309740.204	38				
	Lama_sekolah	2338.511	38				
	Prev_TBC	2.700	38				
	Prev_malaria	11.780	38				
Corrected Total	PDRB	1.143E10	37				
	AHH	365.733	37				
	AMH	1751.175	37				
	Lama_sekolah	88.262	37				
	Prev_TBC	.754	37				
	Prev_malaria	2.676	37				

a. R Squared = ,089 (Adjusted R Squared = ,037)

b. R Squared = ,724 (Adjusted R Squared = ,708)

c. R Squared = ,591 (Adjusted R Squared = ,568)

d. R Squared = ,492 (Adjusted R Squared = ,463)

e. R Squared = ,127 (Adjusted R Squared = ,077)

f. R Squared = ,456 (Adjusted R Squared = ,425)

Between-Subjects SSCP Matrix

			PDRB	AHH	AMH	Lama_sekolah	Prev_TBC	Prev_malaria
Hypothesis	Intercept	PDRB	4.621E9	2.783E7	3.634E7	3065343.881	9.621E4	214578.729
		AHH	2.783E7	1.677E5	2.189E5	18463.411	579.496	1292.467
		AMH	3.634E7	2.189E5	2.858E5	24106.186	756.602	1687.470
		Lama_sekolah	3.065E6	1.846E4	2.411E4	2033.355	63.819	142.338
		Prev_TBC	9.621E4	579.496	756.602	63.819	2.003	4.467
		Prev_malaria	2.146E5	1.292E3	1.687E3	142.338	4.467	9.964
	Klaster	PDRB	1.023E9	2.906E5	4.433E5	45776.523	2868.513	34968.186
		AHH	2.906E5	264.737	517.762	99.864	-3.185	7.793
		AMH	4.433E5	517.762	1.035E3	206.649	-7.360	10.550
		Lama_sekolah	4.578E4	99.864	206.649	43.462	-1.779	.543
		Prev_TBC	2.869E3	-3.185	-7.360	-1.779	.096	.145
		Prev_malaria	3.497E4	7.793	10.550	.543	.145	1.221
Error	PDRB	1.041E10	9.503E4	8.564E5	326718.923	1.164E4	-43160.322	
	AHH	9.503E4	100.996	112.829	27.070	-.802	1.857	
	AMH	8.564E5	112.829	716.356	148.831	3.184	-2.308	
	Lama_sekolah	3.267E5	27.070	148.831	44.800	.761	-1.177	
	Prev_TBC	1.164E4	-.802	3.184	.761	.658	-.375	
	Prev_malaria	4.316E4	1.857	-2.308	-1.177	-.375	1.455	

Based on Type III Sum of Squares

LAMPIRAN H
SURAT PERNYATAAN DATA SEKUNDER BPS

SURAT
PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Jurusan Statistika FMIPA ITS:

Nama : Marina Marsudi Putri

NRP : 1311 100 028

menyatakan bahwa data yang digunakan dalam Tugas Akhir/~~Thesis~~ ini merupakan data sekunder yang diambil dari ~~penelitian/buku/Tugas Akhir/Thesis~~/publikasi lainnya yaitu:

Sumber : www.jatim.bps.go.id

Keterangan : Data IPM dan PDRB Provinsi Jawa Timur

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui
Pembimbing Tugas Akhir

Surabaya, 26 Januari 2015

(Dr. Dra. Kartika Fithriasari, M.Si.)
NIP. 19691212 199303 2 002

(Marina Marsudi Putri)
NRP. 1311 100 028

LAMPIRAN I
SURAT PERNYATAAN DATA SEKUNDER RISKESDAS

SURAT
PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Jurusan Statistika FMIPA ITS:

Nama : Marina Marsudi Putri

NRP : 1311 100 028

menyatakan bahwa data yang digunakan dalam Tugas Akhir/~~Thesis~~ ini merupakan data sekunder yang diambil dari ~~penelitian/buku/Tugas Akhir/Thesis~~/publikasi lainnya yaitu:

Sumber : www.litbang.depkes.go.id

Keterangan : Data Riskesdas Provinsi Jawa Timur

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui
Pembimbing Tugas Akhir

Surabaya, 26 Januari 2015

(Dr. Dra. Kartika Fithriasari, M.Si.)
NIP. 19691212 199303 2 002

(Marina Marsudi Putri)
NRP. 1311 100 028

DAFTAR LAMPIRAN

	halaman
Lampiran A Data Penelitian.....	59
Lampiran B Hasil Analisis Faktor	61
Lampiran C Nilai <i>Score Factor</i>	63
Lampiran D Nilai <i>Pseudo F</i> , <i>Icdrate</i> , SSW, SSB.....	64
Lampiran E Nilai <i>K-Means</i>	70
Lampiran F Program Kohonen SOM di Matlab	71
Lampiran G <i>Output Uji One-way MANOVA</i>	72
Lampiran H Surat Pernyataan Data Sekunder BPS.....	75
Lampiran I Surat Pernyataan Data Sekunder Riskesdas	77

(Halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- Badan Pusat Statistik Provinsi Jawa Timur. (2012). *Tabel Indeks Pembangunan Manusia (IPM) Tahun 2012*. Surabaya: Badan Pusat Statistik Provinsi Jawa Timur.
- Badan Pusat Statistik. (2011). *Kajian Indikator Kesehatan (Laporan Sosial 2010)*. Jakarta: Badan Pusat Statistik.
- Ettaouil, M., Abdelatifi, E., Belhabib, F., Moutaouakil, Karim El. (2012). Learning Algorithm of Kohonen Network with Selection Phase. *WSEAS Transactions on Computers Issue 11, Volume 11, November 2012*. E-ISSN: 2224-2872.
- Fausett, Laurence. (1994). *Fundamental Neural Network: Architectures, Algorithm, and Applications*. New Jersey: Prentice Hall Inc.
- Gerring, J., Thacker, Strom. C., Enikolopov, R., Julian, A., and Maguire, M. (2013). Assesing Health System Performance: A Model-based Approach. *Social Science & Medicine* 93, 21-28.
- Hermawan, Arief. (2006). *Jaringan Syaraf Tiruan Teori dan Aplikasi*. Yogyakarta: Graha ilmu.
- Johnson. R. and Wichern. D. 2007. *Applied Multivariate Statistical Analysis. 6th Edition*. New Jersey: Prentice Hall.
- Kusumadewi, S. (2004). *Membangun Jaringan Syaraf Tiruan (menggunakan MATLAB & Excel Link)*. Yogyakarta: ANDI.
- Mingoti, S.A. & Lima, J.O. (2006). *Comparing SOM Neural Network with Fuzy C-Means, K-Means and Traditional Hierarchical Clustering Algorithm*: 1742-1759.
- Morrison, Donald. F. (1990). *Multivariate Statistical Methods Third Edition*. USA: McGRAW-Hill Book Company.
- Nugroho, Juniarto Setyo. (2014). *Pengelompokan Kabupaten atau Kota di Jawa Timur Menurut Indikator Kesehatan Tahun 2011 dengan Metode Biplot*. Surabaya: Institut Teknologi Sepuluh Nopember.

- Orpin, A. R., & Kostylev, V. E. (2006). *Toward a Statistically Valid Method of Textural Sea Floor Characterization of Benthic Habitats*. *Marine Geology* , 209-222.
- Ramadhani, Anisa. B. C. (2014). *Pemodelan Derajat Kesehatan Kota Surabaya Tahun 2012*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Shovalina, Mei Rizka. (2014). *Analisis Cluster pada Indikator yang Membentuk Indeks Pembangunan Kesehatan Masyarakat di Kabupaten/Kota Provinsi Jawa Timur*. Surabaya: Institut Teknologi Sepuluh Nopember
- Sarle, Warren. S. (1994). Neural Network and Statistical Models. *Proceedings of Nineteenth Annual SAS Users Group International Conference, April, 1994*.
- Siang, Jong Jek. (2009). *Jaringan Syaraf Tiruan & Pemrogramannya Menggunakan MATLAB*. Yogyakarta: ANDI.
- Toor, K. A. and Singh, A. (2013). Analysis of Clustering Algorithm Based on Number of Cluster, Error Rate, Computation Time and Map Topology on Large Data Set. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) Volume 2, Issue 6, November-December 2013*. ISSN 2278-6856.

BIODATA PENULIS



Marina Marsudi Putri lahir di Surabaya pada tanggal 7 April 1994. Penulis yang memiliki hobi menonton film dan membaca novel ini merupakan putri pertama dari Agus Marsudi dan Erwati. Sebelumnya, penulis telah menempuh pendidikan formal di SDN Wonokusumo V Surabaya, SMPN 1 Surabaya dan SMA Negeri 2 Surabaya.

Setelah lulus SMA penulis melanjutkan pendidikan menuju jenjang perguruan tinggi di jurusan Statistika ITS melalui jalur masuk SNMPTN Undangan pada tahun 2011. Pada tahun kedua sebagai mahasiswa di Statistika ITS Penulis aktif berorganisasi sebagai Staf SOSMAS HIMASTA ITS periode 2012-2013. Pada tahun 2013 Penulis berkesempatan untuk menjadi panitia kegiatan STATION yang merupakan lomba Statistika tingkat Nasional yang diadakan oleh HIMASTA ITS. Pada tahun yang sama penulis juga ikut serta dalam panitia kegiatan ITS Expo 2013. Semasa perkuliahan penulis aktif menulis PKM (Program Kreativitas Mahasiswa) dan berkesempatan memperoleh dana hibah PKM dari DIKTI pada tahun 2012, 2014 dan 2015.

Untuk berdiskusi lebih lanjut mengenai tugas akhir, hubungi penulis melalui:

Email: marinamarsudi@rocketmail.com