

ANALISIS KARAKTERISTIK POLA BELANJA KELUARGA DENGAN ANALISIS KLASTER

Nur Arina Hidayati

Program Studi Pendidikan Matematika FKIP UAD

Jl. Prof. Dr. Soepomo, SH. Janturan Yogyakarta

nurarinahidayati@gmail.com

ABSTRAK

Analisis klaster merupakan salah satu alat untuk mengelompokkan sejumlah n obyek berdasarkan p variat yang secara relatif mempunyai kesamaan karakteristik di antara obyek-obyek tersebut. Tujuan utama analisis klaster adalah menempatkan sekumpulan objek ke dalam dua grup atau lebih berdasarkan kesamaan objek atas dasar berbagai karakteristik.

Segmentasi pola belanja merupakan bagian dari segmentasi pasar dalam kategori psikografi. Segmentasi ini mengelompokkan pasar dalam variabel gaya hidup. Dalam hal ini gaya hidup berdasarkan pola belanja.

Kesimpulan dari segmentasi pola belanja dengan studi kasus data rata-rata pengeluaran belanja makanan dari beberapa keluarga di negara Perancis, diperoleh Klaster 1 yang beranggotakan responden dengan tipe keluarga MA dan EM dengan jumlah anak 2 dan 3 memiliki prosentase rata-rata belanja 79% dari seluruh transaksi. Klaster 2 yang beranggotakan responden dengan tipe keluarga MA dan EM dengan jumlah anak 4 dan 5, dan tipe CA dengan jumlah anak 2, memiliki prosentase rata-rata belanja 106%. Klaster 3 yang beranggotakan responden dengan tipe keluarga CA dengan jumlah anak 3, 4 dan 5. Karakter khusus dari keluarga ini adalah mempunyai minat belanja yang tinggi dibanding klaster yang lain, tampak dari prosentase rata-rata belanja 118% dari seluruh transaksi

Kata kunci : Analisis klaster, segmentasi pasar, karakteristik

ABSTRACT

Cluster analysis is a tool for classifying n objects based on p variat relatively have a common characteristic among these objects. The main goal is to put a set of cluster analysis of objects into two or more groups based on similarity of objects on the basis of various characteristics.

Segmentation pattern of spending is part of the category of psychographic segmentation. This segmentation variable grouping in the lifestyle market. In this case a lifestyle based on spending patterns.

Conclusion of the segmentation pattern of shopping with a case study data is the average food expenditure of a few families in the country of France, obtained Cluster 1 is composed of respondents with family type MA and EM with child number 2 and 3 have an average percentage of spending 79% of all transaction. Cluster 2 is composed of respondents with family type MA and EM with child number 4 and 5, and the type of CA with child number 2, had an average percentage of spending 106%. Cluster 3 is composed of respondents with family-type CA with child number 3, 4 and 5. The special character of this family have an interest expenditure is higher than the other clusters, it appears from the average percentage of spending 118% of all transactions

Keywords : Cluster analysis, Segmentation pattern, Characteristic

Pendahuluan

Analisis kluster merupakan salah satu teknik multivariate metode interdependensi (saling ketergantungan). Oleh karena itu, dalam analisis kluster tidak ada perbedaan antara variabel bebas (*independent variabel*) dan variabel terikat (*dependent variabel*).

Analisis kluster digunakan untuk mengelompokkan data observasi yang hanya berdasarkan pada informasi yang ditemukan dalam data, dimana data tersebut harus menggambarkan observasi dan hubungannya.

Analisis kluster adalah teknik yang digunakan untuk menggabungkan observasi ke dalam kelompok kluster atau kluster (Sharma, 1996:185), sedemikian sehingga:

1. Setiap kelompok atau kluster homogen mempunyai karakteristik tertentu. Hal ini berarti bahwa observasi dalam setiap kelompok sama dengan observasi lain dalam satu kelompok yang sama;
2. Setiap kelompok seharusnya berbeda dari kelompok lain dengan karakteristik yang sama. Hal ini berarti bahwa observasi dalam kelompok yang satu seharusnya berbeda dari observasi dalam kelompok lain.

Tujuan utama analisis kluster adalah menempatkan sekumpulan objek ke dalam dua grup atau lebih berdasarkan kesamaan objek atas dasar berbagai karakteristik. Hasil kluster suatu obyek harus memiliki internal (*within cluster*) homogenitas yang tinggi dan memiliki eksternal (*between cluster*) heterogenitas yang tinggi. Semakin besar kemiripan (homogenitas) dalam kelompok dan perbedaan antar kelompok maka klustering akan lebih baik atau lebih berbeda (Tan et al,2006:490).

Dalam analisis kluster, pengelompokan ke dalam kluster dilakukan dengan menggunakan teknik-teknik yang berawal dari kemiripan antar semua pasangan observasi. Kemiripan ini didasarkan pada berapa ukuran jarak antara observasi satu dengan observasi yang lain.

Pada prinsipnya analisis kluster merupakan proses untuk mereduksi sejumlah objek yang besar menjadi lebih sedikit yang disebut kluster. Analisis kluster disebut juga *Q-analysis*, *classification analysis*, pengenalan pola (*pattern recognition*), analisis segmentasi (*numerical taxonomy*).

Segmentasi pasar

Segmentasi pasar adalah proses pemisahan sebuah pasar menjadi

kelompok-kelompok pembeli atau produk yang sejenis dan memiliki kelompok (atau kelompok-kelompok) yang paling tepat untuk dilayani perusahaan (Peter, 2000). Hal ini merupakan upaya paling efektif dalam menyingkapi peluang pasar. Beberapa jenis segmentasi utama adalah geografi, demografi, psikografi, perilaku dan manfaat (Morissan, 2009). Secara garis besar penjelasan masing-masing jenis sebagai berikut:

1. Segmentasi Geografi, merupakan pembagian pasar menjadi unit-unit geografis yang berbeda. Misalnya: wilayah, negara bagian, propinsi, kota dan kepulauan.
2. Segmentasi Demografi, dimana pasar dikelompokkan berdasarkan variabel-variabel pendapatan, jenis kelamin, pendidikan, jumlah penduduk, usia, ukuran keluarga, pekerjaan, agama, ras, generasi, kewarganegaraan dan kelas sosial.
3. Segmentasi Psikografi, mengelompokkan pasar dalam variabel gaya hidup, nilai dan kepribadian. Gaya hidup juga ditunjukkan oleh orang-orang yang menonjol pada kelas sosial. Minat terhadap suatu produk juga dipengaruhi oleh gaya hidup.
4. Segmentasi Perilaku, membagi kelompok berdasarkan status pemakai, kejadian, tingkat

penggunaan, status kesetiaan, tahap kesiapan pembeli, dan sikap. Pasar disini dapat dikelompokkan menjadi bukan pemakai, bekas pemakai, pemakai potensial, pemakai pertama kali, dan pemakai tetap dari suatu produk.

5. Segmentasi Manfaat

Mengklasifikasi pasar berdasarkan atribut/nilai atau manfaat yang terkandung dalam suatu produk.

Pada tulisan ini digunakan segmentasi psikografi, dimana objek dikelompokkan berdasarkan produk-produk bahan olahan makanan yang dibeli dalam sebuah supermarket. Berdasarkan pola belanja terhadap produk-produk tersebut, akan diketahui pola hidup objek-objek tersebut.

Metode Pengelompokan (*Clustering*)

Secara umum metode pengelompokan dalam analisis kluster dibedakan menjadi metode hirarki (*Hierarchical Clustering Methode*) dan metode non hirarki (*Nonhierarchical Clustering Methode*).

A. Metode Hirarki

Pada dasarnya metode ini dibedakan menjadi dua metode pengelompokan, yaitu:

1. Metode Agglomeratif

Proses pengelompokan dengan pendekatan aglomeratif (*Down to Top*) dimulai dengan n kluster sehingga masing-masing kluster memiliki tepat satu objek, kemudian menentukan dua kluster terdekat dan menggabungkan dua kluster tersebut menjadi satu kluster baru. Proses penggabungan dua kluster ini diulangi sampai diperoleh satu kluster yang memuat semua himpunan data. Hasil analisis kluster dari metode ini dapat disajikan dalam bentuk dendogram.

Dalam metode Agglomeratif ini terdapat lima metode yang cukup terkenal, yaitu:

a. *Single Linkage*

Prosedur ini didasarkan pada jarak terkecil. Jika dua obyek terpisah oleh jarak yang pendek maka kedua obyek tersebut akan akan digabung menjadi satu kluster, demikian seterusnya.

b. *Complete Linkage*

Berlawanan dengan *Single Linkage*, prosedur ini pengelompokannya berdasarkan jarak terjauh.

c. *Average Linkage*

Prosedur ini hampir sama dengan *Single Linkage* dan

Complete Linkage, namun criteria yang digunakan adalah rata-rata jarak seluruh individu dalam suatu kluster dengan jarak seluruh individu dalam kluster yang lain.

d. *Ward's Methode*

Jarak antara dua kluster dalam metode ini berdasarkan total *sum of square* dua kluster pada masing-masing variabel.

e. *Centroid method*

Jarak dua kluster dalam metode ini berdasarkan jarak centroid dua kluster yang bersangkutan.

2. Metode Divisif

Proses pengelompokan dengan pendekatan metode divisive (*Top to Down*) dimulai dengan n objek yang dikelompokkan menjadi satu kluster, kemudian kluster tersebut dipartisi ke dalam dua kluster pada setiap langkah sampai diperoleh n kluster dengan setiap kluster memiliki satu objek.

B. Metode Non-Hirarki

Dalam metode ini data dibagi dalam k partisi, setiap partisi mewakili sebuah kluster. Secara umum proses metode non-hirarki sebagai berikut:

1. Pilih k centroid kluster awal atau *seed*, dimana k merupakan jumlah kluster yang diinginkan

2. Tempatkan setiap observasi ke dalam klaster yang terdekat
3. Tempatkan kembali setiap observasi ke dalam k klaster menurut aturan penghentian yang sudah di tentukan
4. Proses berhenti jika tidak ada observasi yang berpindah lagi, jika belum ulangi langkah kedua.

Langkah-langkah melakukan analisis klaster

Dalam tulisan ini, digunakan metode pengklasteran hirarki metode agglomeratif. Adapun langkah-langkahnya adalah sebagai berikut:

1. Langkah ini dimulai dengan n klaster sehingga masing-masing klaster memiliki tepat satu objek.
2. Menentukan dua klaster terdekat dan menggabungkan dua klaster tersebut menjadi satu klaster baru. Untuk menentukan jarak terdekat antara dua cluster digunakan *metode ward* atau *varians method*. Metode ini bertujuan memperoleh klaster

yang memiliki varians interval klaster yang sekecil mungkin, yaitu dengan mengkuadratkan jarak Euclid (*Squared Euclidean Distanc*e) antar klaster.

3. Ulangi langkah ke-2 sehingga berakhir sampai diperoleh satu klaster yang memuat semua himpunan data. Data Out put dalam SPSS berupa dendogram.
4. Lakukan profilisasi dan interpretasi klaster. Interpretasi dilakkan untuk mencari karakteristik tiap klaster yang khas.

Olah Data

Dalam tulisan ini, data yang digunakan adalah data rata-rata pengeluaran belanja makanan dari beberapa keluarga di negara Perancis dengan tipe keluarga yang berbeda (tipe MA, tipe EM, tipe CA) dengan jumlah anak 2,3, 4 atau 5. Data ini diambil dari Lebart, Morineau dan Fenelon (1982) yang disajikan dalam tabel berikut ini :

Tabel 1. Daftar Belanja Keluarga di Perancis

No	Type	Bread	Vegetables	Fruits	Meat	Poultry	Milk	Wine
1	MA2	332	428	354	1437	526	247	427
2	EM2	293	559	388	1527	567	239	258
3	CA2	372	767	562	1948	927	235	433
4	MA3	406	563	341	1507	544	324	407
5	EM3	386	608	396	1501	558	319	363
6	CA3	438	843	689	2345	1148	243	341
7	MA4	534	660	367	1620	638	414	407

8	EM4	460	699	484	1856	762	400	416
9	CA4	385	789	621	2366	1149	304	282
10	MA5	655	776	423	1848	759	495	486
11	EM5	584	995	548	2056	893	518	319
12	CA5	515	1097	887	2630	1167	561	284
	Rata - rata	446.7	732.0	505.0	1886.8	803.2	358.3	368.6

Prosedur olah data dilakukan dengan SPSS 15.00, dengan Out Put sebagai berikut:

Cluster

[DataSet1] E:\S2 Matematika UGM\UJIAN Data Mining\French Food Data.sav

Case Processing Summary^{a,b}

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
12	100.0	0	.0	12	100.0

a. Squared Euclidean Distance used

b. Ward Linkage

Ward Linkage

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	4	5	3821.500	0	0	3
2	6	9	12818.000	0	0	9
3	2	4	36425.833	0	1	4
4	1	2	65855.250	0	3	10
5	8	10	96691.750	0	0	6
6	7	8	153977.917	0	5	8
7	3	11	255492.417	0	0	8
8	3	7	430601.250	7	6	10
9	6	12	632336.083	2	0	11
10	1	3	1334487.56	4	8	11
11	1	6	3435249.75	10	9	0

Berdasarkan Tabel *Agglomeration Schedule* diatas diperoleh keterangan sebagai berikut :

Pada **tahap pertama** obyek yang dipasangkan adalah responden **4** dan **5** karena pasangan ini memiliki koefisien jarak paling rendah, dimana pada saat itu objek-objek yang lain sama sekali belum memperoleh pasangan. Oleh karena itu, pada tahap ini jumlah kluster ada **11**. Pada baris pertama ini, pada kolom *next page* tertulis **3**. Artinya pada tahap ke-**2**, pasangan **4** dan **5** akan mendapat anggota baru, yaitu responden yang ke-**4**.

Tahap ke dua, responden **6** dan **9** dipasangkan. Artinya, koefisien jarak terendah yang kedua adalah pasangan iresponden ini. Kemudian pada kolom *next page* tertulis **9**, artinya pada tahap ke-**8**, pasangan **6** dan **9** akan mendapat anggota baru. Pada tahap ini jumlah kluster ada **10**.

Demikian seterusnya, pada setiap tahap jumlah kluster berkurang satu, sehingga pada tahap tahap yang terakhir yaitu ke-**11**, semua responden menjadi satu kluster

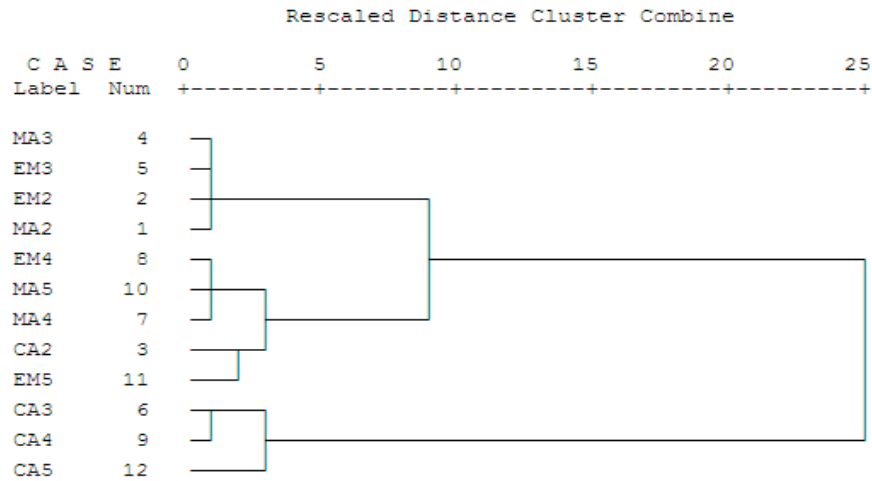
Vertical Icicle

Number of clusters	Case																						
	12:CA5		9:CA4		6:CA3		10:MA5		8:EM4		7:MA4		11:EM5		3:CA2		5:EM3		4:MA3		2:EM2		1:MA2
1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
5	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
6	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
7	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
8	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
9	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
10	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
11	X		X		X		X		X		X		X		X		X		X		X		X

Dendrogram

***** H I E R A R C H I C A L C L U S T E R A N A L Y S I S *****

Dendrogram using Ward Method



Berkaitan dengan berapa banyaknya kluster atau jumlah kluster yang dapat terbentuk dari kasus tersebut, dendrogram diatas dapat dijadikan acuan untuk menentukan jumlah

cluster. Dari dendogram di atas dapat terlihat jumlah kluster yang dapat dibentuk adalah **tiga kluster**, dengan anggota sebagai berikut :

Tabel 2. Pengelompokan Keluarga Berdasarkan Klaster

Klaster	Case	Klaster	Case	Klaster	Case
Klaster 1	4 : MA3	Klaster 2	6 : EM4	Klaster 3	6 : CA3
	5 : EM3		10 : MA5		9 : CA4
	2 : EM2		7 : MA4		12 : CA5
	1 : MA2		3 : CA2		
			11 : EM5		

Tahap Selanjutnya Adalah Interpretasi dan Pembuatan Profil.

Untuk menginterpretasikan kluster dan membuat profil,

digunakan rata-rata setiap kluster pada setiap variable (centroid).

Berikut tabel Profil dari masing-masing Klaster :

Tabel 3. Profil Kluster Satu

No	type	Bread	Vegetables	Fruits	Meat	Poultry	Milk	Wine
4	MA3	406	563	341	1507	544	324	407
5	EM3	386	608	396	1501	558	319	363
2	EM2	293	559	388	1527	567	239	258
1	MA2	332	428	354	1437	526	247	427
Rata - rata		354.25	539.5	369.7 5	1493	548.75	282.2 5	363.7 5
Prosentase Belanja		79%	73%	73%	79%	68%	79%	99%

Tabel 4. Profil Kluster Dua

No	type	Bread	Vegetables	Fruits	Meat	Poultry	Milk	Wine
8	EM4	460	699	484	1856	762	400	416
10	MA5	655	776	423	1848	759	495	486
7	MA4	534	660	367	1620	638	414	407
3	CA2	372	767	562	1948	927	235	433
11	EM5	584	995	548	2056	893	518	319
Rata - rata		521	779.4	476.8	1865.6	795.8	412.4	412.2
Prosentase Belanja		117%	106%	94%	99%	99%	115%	112%

Tabel 5. Profil Kluster Tiga

No	type	Bread	Vegetables	Fruits	Meat	Poultry	Milk	Wine
6	CA3	438	843	689	2345	1148	243	341
9	CA4	385	789	621	2366	1149	304	282
12	CA5	515	1097	887	2630	1167	561	284
Rata - rata		446.00	909.67	732.33	2447.00	1154.67	369.33	302.33
Prosentase Belanja		100%	123%	145%	130%	144%	103%	82%

kluster, disajikan rekapitulasi prosentase belanja dari masing-masing kluster sebagai berikut Untuk

memudahkan dalam mengamati karakteristik pola belanja masing-masing:

Tabel 6. Prosentase Pola Belanja

	Prosentase Pola Belanja Klaster 1	Prosentase Pola Belanja Klaster 2	Prosentase Pola Belanja Klaster 3
Bread	79%	117%	100%
Vegetables	73%	106%	123%
Fruits	73%	94%	145%
Meat	79%	99%	130%
Poultry	68%	99%	144%
Milk	79%	115%	103%
Wine	99%	112%	82%
Rata-rata	79%	106%	118%

Kesimpulan

Analisa tabel Prosentase Pola Belanja:

1. **Klaster 1** yang beranggotakan responden dengan **tipe keluarga MA dan EM** dengan jumlah **anak 2 dan 3** memiliki prosentase rata-rata belanja **79%** dari seluruh transaksi. Dapat dilihat prosentase belanja keluarga ini paling kecil dibandingkan tipe keluarga yang lain. Tipe keluarga ini jarang mengkonsumsi olahan daging itik, terlihat dari prosentase yang kecil jika dibandingkan dengan item yang lain, yaitu sebesar 68%. Satu hal yang menarik dari tipe keluarga ini adalah keluarga ini **mempunyai karakter suka mengkonsumsi wine**, tampak dari prosentase yang paling tinggi dibandingkan dengan item yang lain yaitu **99%**, selain itu juga jika dibandingkan dengan klaster yang lain konsumsi wine tipe keluarga ini juga cukup tinggi bahkan melebihi klaster yang ke 3 walaupun klaster yang ketiga prosentase belanjanya lebih besar.
2. **Klaster 2** yang beranggotakan responden dengan **tipe keluarga MA dan EM** dengan jumlah **anak 4 dan 5**, dan **tipe CA** dengan jumlah **anak 2**, memiliki prosentase rata-rata belanja **106%** dari seluruh transaksi. Keluarga ini **nampaknya mempunyai kebiasaan sarapan dengan bread dan milk**, dilihat dari prosentase pembelian yang cukup tinggi jika dibandingkan dengan item yang lain dan klaster yang lain, yaitu prosentase **bread sebesar 117%** dan prosentase **milk 115%**.
3. **Klaster 3** yang beranggotakan responden dengan **tipe keluarga CA** dengan jumlah **anak 3, 4 dan 5**. Karakter khusus dari keluarga ini adalah mempunyai sense belanja yang tinggi dibanding klaster yang

lain, tampak dari prosentase rata-rata belanja **118%** dari seluruh transaksi. Akan tetapi tipe keluarga ini **cukup memperhatikan kesehatan**, dapat dilihat dari prosentase belanja buah yang sangat tinggi baik dibandingkan dengan item yang lain maupun klaster tipe keluarga yang lain (**145%**), selain itu pula konsumsi wine dari tipe keluarga ini paling kecil jika dibandingkan dengan yang lain (**82%**). Keluarga ini juga sangat menyukai olahan daging itik (**144%**).

Pustaka

- Ghozali, I., *Aplikasi Multivariat dengan Program SPSS*, Badan Penerbit U ndip, Semarang, 2001
- Morissan, *Sales-Marketing Pokok Bahasan: Segmentasi Pasar*, 2009, URL:
<http://pksm.mercubuana.ac.id/new/elearning/files/modul/41003-6-256559724556.pdf>
- Peter, J. P and Olson, J.C., *Consumer Behavior : Perilaku Konsumen dan Strategi Pemasaran*, Erlangga, Jakarta, 2000.
- Rosiatun, A., Widiharih,T., Safitri, D., Analisis Klaster Untuk Segmentasi Pemirsas Program Berita Sore Stasiun TV Swasta, FMIPA Universitas Diponegoro.
- Sharma, S., *Applied Multivariate Techniques*, John Willey & Sons, New York, 1996.
- Simamora, B., *Analisis Multivariat Pemasaran*, Gramedia Pustaka Utama, Jakarta, 2005.

Tan, P., Steinbach, M., & Kumar, V., (2006). *Introduction to Data Mining*, Pearson Education.

