

ANALISIS KLASTER PENGELOMPOKKAN KOTA PADA DATA PENJUALAN SONGKOK NASIONAL UMKM KABUPATEN GRESIK

Anik Rufaidah

Program Studi Teknik Industri Fakultas Teknik Universitas Qomaruddin

Jalan Raya No. 01 Bungah Gresik 61152 Indonesia

Anikrufaidah99@gmail.com

ABSTRAK

Kabupaten Gresik merupakan kota produksi songkok Nasional, yang mana kota tersebut banyak UMKM memproduksi songkok tersebut. Dan pemesanan songkok tersebut tersebar beberapa kota yang di Indonesia, yang mana pemesanan songkok tersebut sangat fluktuatif. Sehingga pihak pengerajin ingin mengetahui pengelompokan kota-kota yang terbagi secara signifikan dengan kesamaan yang sama, dengan menggunakan metode analisis kluster dapat dilihat pengelompokan kota-kota yang mempunyai karakteristik pemesanan yang sama. Dengan menggunakan analisis kluster dengan metode *Hierarchical Cluster* dan metode *K-means Cluster* dengan bantuan software SPSS didapat pengelompokan 2 (dua) kluster.

Kata Kunci: *Kluster Pengelompokan Kota dan Umkm*

1. PENDAHULUAN

Kabupaten Gresik merupakan kota penghasil songkok nasional, yang mana di kota tersebut banyak UMKM yang memproduksi songkok nasional tersebut. Dimana penjualan songkok nasional terjual kebeberapa kota di Indonesia, dari penjualan di beberapa kota tersebut permintaan setiap kota yang tidak menentu. Sehingga para pengerajin ingin mengetahui daerah-daerah/ kota-kota mana saja yang mempunyai prioritas untuk memenuhi kebutuhan pemesanan songkok nasional tersebut.

Analisis kluster dapat digunakan untuk mengelompokkan kota-kota pemesanan penjualan songkok nasional tersebut. Dari hasil analisis kluster tersebut dapat dilihat pengelompokannya berdasarkan kesamaan antar kelompok dari sisi banyaknya penjualan yang dilakukan. Sehingga hasil analisis kluster tersebut dapat dilihat kota-kota yang mana saja yang harus

diprioritaskan dalam waktu pemesanan yang sama. Dengan demikian proses produksi yang dilakukan dapat terfokus untuk daerah-daerah yang punya keragaman yang sama.

Dengan adanya tidak menentu jumlah order pada tiap-tiap kota, sehingga diharapkan analisis klaster ini dapat melihat bagaimana pengelompokkan tiap-tiap kota yang ada. Dengan adanya pengelompokkan pada kota-kota yang ada dapat dilakukan mana kota-kota yang memerlukan order yang banyak, selain itu dapat dilakukan metode marketing yang sesuai sehingga diharapkan adanya pemerataan terhadap kota-kota yang mengalami pemesanan kecil.

Pengelompokkan ini dapat dilakukan dengan menggunakan metode *Hierarchical Cluster* dan metode *K-means Cluster* yang mana kedua metode ini dapat menghasilkan kelompok kota-kota yang dalam satu kelompok mempunyai kemiripan yang besar dan di kelompok lain mempunyai kemiripan yang besar antar kelompoknya. Diharapkan dengan pengelompok ini pengerajin dapat melihat pemesanan mana yang harus didahulukan pada waktu adanya lonjakan pemesanan semua konsumen.

2. TINJAUAN PUSTAKA

Analisis *cluster* adalah analisis statistika yang bertujuan untuk mengelompokkan data sedemikian sehingga data yang berada dalam kelompok yang sama mempunyai sifat yang relatif homogen daripada data yang berada dalam kelompok yang berbeda. Ditinjau dari hal-hal yang dikelompokkan, analisis *cluster* dibagi menjadi dua macam, yaitu pengelompokan observasi dan pengelompokan variabel. Dalam pembahasan ini, pengelompokan yang dilakukan adalah pengelompokan observasi.

Jika ditinjau dari metode pengelompokannya, analisis *cluster* memiliki dua metode, yaitu:

1. Metode *Hierarki*.

Metode ini digunakan untuk mencari struktur pengelompokan dari objek-objek. Jadi, hasil pengelompokannya disajikan secara *hierarki* atau berjenjang. Metode *hierarki* ini terdiri dari dua cara, yaitu:

a). Penggabungan (*Agglomerative*).

Cara ini digunakan jika masing-masing objek dianggap satu kelompok kemudian antar kelompok yang jaraknya berdekatan bergabung menjadi satu kelompok.

b) Pemecahan (*Devise*).

Cara ini digunakan jika pada awalnya semua objek berada dalam satu kelompok. Setelah itu, sifat paling beda dipisahkan dan membentuk satu kelompok yang lain. Proses tersebut berlanjut sampai semua objek tersebut masing-masing membentuk satu kelompok.

2. Metode Tidak Hierarki

Metode ini digunakan apabila jumlah kelompok yang diinginkan sudah diketahui dan biasanya dipakai untuk mengelompokkan data yang ukurannya besar. Dalam proses penggabungan kelompok dengan metode hierarki selalu diikuti dengan perbaikan matriks jarak. Suatu fungsi disebut jarak jika mempunyai sifat tak negatif ($d_{ij} \geq 0$) dan ($d_{ij} = 0$) jika $i = j$ simetris ($d_{ij} = d_{ji}$), panjang salah satu sisi segitiga selalu lebih kecil atau sama dengan jumlah dua sisi yang lain ($d_{ij} \leq d_{ik} + d_{jk}$). Beberapa macam jarak yang biasa dipakai di dalam analisis kelompok:

Jarak *Euclidean*:

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

Jarak *Manhattan*:

$$d_{ij} = \sum_{k=1}^p |x_{ik} - x_{jk}|$$

Jarak *Pearson*:

$$d_{ij} = \sqrt{\sum_{k=1}^p \frac{(x_{ik} - x_{jk})^2}{\text{var}(x_k)}}$$

Jarak Korelasi:

$$d_{ij} = 1 - r_{ij}$$

Jarak Korelasi Mutlak:

$$d_{ij} = 1 - |r_{ij}|$$

Dalam banyak penelitian, jarak *euclidean* adalah jarak yang paling sering digunakan, begitu juga dengan kuadrat jarak *euclidean*.

Metode-metode pengelompokan *hierarki* dibedakan berdasarkan konsep jarak antar kelompok, penentuan jarak antar kelompok untuk metode-metode tersebut adalah sebagai berikut.

1. Metode *Single Linkage*

Metode ini mengelompokkan dua objek yang mempunyai jarak terdekat terlebih dahulu.

Jarak antar kelompok (*i,j*) dengan *k* adalah :

$$d(i,j)k = \min(d_{ik}, d_{jk})$$

2. Metode *Complete Linkage*

Metode ini akan mengelompokkan dua objek yang mempunyai jarak terjauh terlebih dahulu. Jarak antar kelompok (*i,j*) dengan *k* adalah :

$$d(i,j)k = \max(d_{ik}, d_{jk})$$

3. Metode *Average Linkage*

Metode ini mengelompokkan objek berdasarkan jarak rata-rata yang didapat dengan melakukan rata-rata semua jarak objek terlebih dahulu. Jarak antar kelompok (*i,j*) dengan *k* adalah :

$$d(i,j)k = \text{average}(d_{ik}, d_{jk})$$

4. Metode *Median Linkage*

Pada metode ini, jarak antara dua *cluster* adalah jarak di antara *centroid cluster* tersebut. *Centroid* adalah rata-rata jarak yang ada pada sebuah *cluster* yang didapat dengan melakukan rata-rata pada semua anggota suatu *cluster* tertentu. Dengan metode ini, setiap terjadi *cluster* baru, akan terjadi perhitungan ulang *centroid* hingga terbentuk *cluster* tetap. Jarak antar kelompok (*i,j*) dengan *k* adalah :

$$d(i,j)k = \text{median}(d_{ik}, d_{jk})$$

3. MOTODE PENELITIAN

3.1. Sumber Data dan Variabel Penelitian

Pada penelitian ini dilakukan pengamatan untuk penjualan pada 2 tahun terakhir yaitu pada data penjualan tahun 2017 dan 2018 yang dihitung banyaknya penjualan perbulannya dengan hasil. Dengan penjualan yang disorder oleh beberapa kota yang ada di Indonesia diantaranya:

1. Kota Babat Lamongan
2. Kota Serang Rembang
3. Kota Mojokerto
4. Kota Tumenggung
5. Kota Purwokerto
6. Kota Bojonegoro
7. Kota Tuban
8. Kota Kebumen
9. Kota Cilacap
10. Kota Nganjuk
11. Kota Magelang

Penjualan dari Songkok Nasional yang bearsal dari UMKM kabupaten Gresik terutama pada UD. Ikkal mengahsilakn pola keragaman yang berbeda, sehingga pengeajin ingin mengetahui katategori atau kelompok-kelompok customer yang mempunyai kesamaan yang dilihat dari tingkat ordernya. Sehingga dari penjualan dibeberapa kota tersebut dilakukan analisis klaster dengan metode *Hirarchical Cluster* dan metode *K-means Cluster*.

Metode Analisis

Pada analisis klaster yang dilakukan pada awalnya adalah mendefinisikan variabel-variabel yang akan digunakan pada analisis klaster tersebut, setelah itu menentukan jarak yang tepat untuk dipilih. Yang mana penentuan jarak tersebut dapat menentukan kemiripan dari objek yang dikelompokkan tersebut. Selanjutnya melakukan tahapan klaster dengan menggunakan metode yang dilakukan. Pemilihan metode yang terbaik yang mana dalam setiap klaster mempunyai anggota yang relative homogen.

Yang dilakukan dalam analisis terakhir adalah menentukan keanggotaan dari masing-masing kluster yang terpilih tersebut. Diharapkan dalam setiap kluster mempunyai nilai keragaman yang kecil dari pada keragaman antar kluster yang terbentuk tersebut. Sehingga anggota yang tidak mirip akan dikelompokkan dalam kelompok yang berbeda yang mempunyai kemiripan dengan yang lain.

Dalam pengklasteran tujuan yang dilakukan adalah untuk menentukan obyek yang mempunyai kemiripan yang sama dalam satu kluster dan mempunyai kemiripan yang berbeda antar kluster. Dimana pendekatan yang bisa dilakukan untuk mengukur kemiripan diantaranya jarak (*distance*) antar obyek.

Cara yang dilakukan dalam pengukuran jarak, yaitu:

- a. Dengan menggunakan jarak, yaitu jarak berupa akar kuadrat perbedaan nilai untuk tiap variabel.

Jika $x = (x_1, x_2, \dots, x_p)$ dan $y = (y_1, y_2, \dots, y_p)$ maka

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2}$$

- b. Menggunakan jarak kuadrat *euclidiean* (*square euclidean distance*)
- c. Diantara dua objek merupakan jumlah nilai perbedaan mutlak untuk tiap variabel (*The City Block or Manhattan Distance*). Jarak ini juga disebut jarak *Minkowski*.

Jika $\bar{x}' = (x_1, x_2, \dots, x_p)$; p adalah variabel. Maka $x = (x_{i1}, x_{i2}, \dots, x_{ip})$ adalah kumpulan variabel pada obyek ke-i $d_{ij} = [\sum_{k=1}^p [|x_{ik} - y_{ik}|^r]]^{1/r}$

Dengan d_{ij} = jarak antara dua obyek ke-I dan obyek ke-j.

- d. *The Chebyshev Distance* antar dua obyek yang merupakan jumlahan nilai pada perbedaan mutlak dari setiap variabel.

Dimana variabel X_1, X_2, \dots, X_p dari setiap pengukuran dilakukan terhadap setiap individu anggota. misalkan n adalah individu dari a_1, a_2, \dots, a_n . Hasil pengukuran variabel X_j diaman individu a_i dinyatakan dengan d_{ij} . Sehingga:

$$d_{ij} = \begin{bmatrix} d_{11} & \dots & d_{1p} \\ \dots & \dots & \dots \\ d_{n1} & \dots & d_{np} \end{bmatrix}$$

Dimana d_{ij} adalah jarak antara individu ke-I dan ke-j, dengan anggota n dan p yang sama.

Untuk jarak (a,b) mempunyai sifat:

- a. Simetri, $d(a,b) = d(b,a)$
- b. Positif, $d(a,b) \geq 0$
- c. $d(a_1, a_2) = 0$
- d. (a,b) meningkat seiring tidak miripnya a dan b
- e. $d(a,c) \leq d(a,b) + d(b,c)$

Yang mana semakin kecil d , maka semakin besar kemiripan antar pengamatan tersebut. Dan sebaliknya semakin besar d , maka semakin besar kemiripan antar pengamatan tersebut. Dengan menggunakan jarak Euclidian yaitu jarak antar individu pada p variabel maka jarak individu ke- i dan ke- j .

Pada penelitian ini data diolah menggunakan *software* SPSS dengan tahapan analisis data yang digunakan adalah sebagai berikut:

1. Melakukan pengumpulan data penjualan songkok nasional UMKM kabupaten Gresik
2. Melihat nilai diskriptif dari data tersebut
3. Melakukan pengujian multikolinieritas, yang digunakan untuk mengetahui ada atau tidaknya variabel independen yang memiliki kemiripan dengan variabel independen lainnya dengan melihat nilai *Variances Inflation Factor* (VIF).
4. Melakukan pengukuran jarak, dengan menggunakan jarak *Squared Euclidean*.
5. Memilih metode analisis kluster dengan metode *Hierarchical Cluster* dan metode *K-means Cluster*
6. Menentukan banyaknya kluster
7. Melakukan pengecekan pembagian kluster dengan menggunakan *R-Square* (R^2)
8. Melakukan interpretasi hasil analisis kluster tentang kluster yang terbentuk.

4. HASIL PENELITIAN DAN PEMBAHASAN

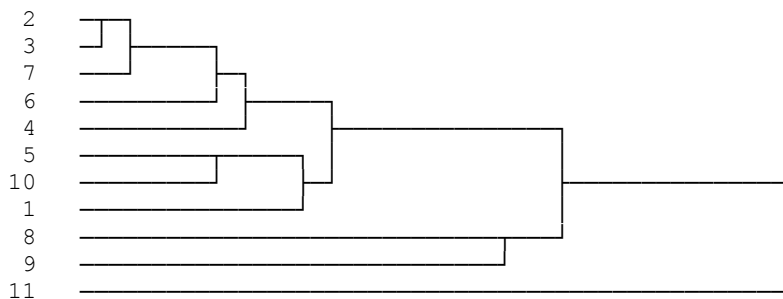
Pembahasan yang dapat dilakukan dari analisis data penjualan songkok nasional pada UMKM di kabupaten Gresik pada 2 tahun terakhir 2017 dan 2018 terhadap 11 kota yang menjadi konsumen dari songkok nasional tersebut diantaranya kota Babat, Serang, Mojokerto, Tumenggung, Purwokerto, Bojonegoro, Tuban, Kebumen, Cilacap, Nganjuk, dan Magelang. Yang mana dari data tersebut dapat dilihat nilai rata-rata dan nilai penyebarannya, hasil dari

data tersebut dapat dilihat pada Tabel 1. Dari data diatas dapat dilihat bahwa nilai rata-rata daerah pemesanan yang paling banyak adalah daerah kebumen, magelang dan cilacap.

Tabel 1. Analisis Deskriptif

NO	Kota	Mean	Sd.Deviasi
1	Babat	37,92	36,62
2	Serang	1,604	4,679
3	Mojokerto	3.56	9.24
4	Tumenggung	19.94	24.26
5	Purwokerto	21.04	35.57
6	Bojonegoro	13.96	26.27
7	Tuban	10.50	17.93
8	Kebumen	50.25	38.55
9	Cilacap	43.5	55.9
10	Nganjuk	15.71	28.50
11	Magelang	43.0	51.5

Untuk analisis kluster metode *Hirarchical Cluster* dari data penjualan songkok Nasional UMKM kabupaten Gresik dengan menggunakan software SPSS dapat dilihat pada gambar 1.



Gambar 1. Analisis Kluster metode *Hirarchical Cluster*

Pada hasil analisis kluster dengan metode *Hirarchical Cluster* menggunakan SPSS dapat dilihat bahwa pengklasteran dengan menggunakan dua kluster didapat hasil sebagai berikut kluster 1 terdiri dari beberapa kota diantaranya Serang, Mojokerto, Tuban, Bojonegoro, Tumenggung, Purwokerto, Nganjuk, Babat. Sedangkan untuk kluster 2 terdiri dari kota Kebumen, Cilacap, dan Magelang.

Dari pengklasteran diatas dapat dilihat untuk kelompok klaster 2 mengalami tingkat order tinggi di bandingkan dengan tingkat order klaster 1. Sehingga penggunaan strategi marketing baru yang harus dilakukan adalah pada klaster 1, dengan harapan adanya srategi marketing yang baru diharapkan dapat meningkatkan jumlah order pada konsumen daerah/ kota yang berada pada klaster 1 diantaranya kota Serang, Mojokerto, Tuban, Bojonegoro, Tumenggung, Purwokerto, Nganjuk, dan Babat.

Sedangkan untuk klaster 2 yang hanya terdiri dari tiga kota diantaranya kota kebumen, Cilacap, dan Magelang. Diharapkan pada klaster 2 maka anggota yang ada dapat bertambah sehingga pemesanan songkok nasional selalu ada kenaikan sehingga UMKM Songkok Nasional di Kabupaten Gresik terus berkembang. Selainin itu diharapkan mengalami perluasn penjualan di seluruh Indonesia sehingga tingkat penjualan yang ada semakin meningkat. Sehingga itu untuk saran yang diharapkan adanya strategi marketing yang sesuai untuk meningkatkan tingkat penjualan songkok nasional tersebut.

Sedangkan untuk analisis klaster dengan metode *K-means Cluster* didapat hasil dengan pengelompokan dua klaster dengan hasil yang didapat dapat dilihat pada tabel 2. Pada tabel 2 tersebut dapat dengan menggunakan perhitungan software SPSS dengan metode *K-means Cluster* sehingga didapat pengelompokan untuk klaster 1 adalah kota Babat, Serang, Mojokerto, Purwokerto, Tuban, Kebumen, Cilacap, dan Nganjuk. Sedangkan untuk klaster 2 terdiri dari kota Tumenggung, Bojonegoro, Magelang. Dari hasil tersebut dapat dilihat bahwa pada klaster yang sama mempunyai kemiripan keragaman, sedangkan antar klaster mempunyai keragaman yang berbeda. Jika melihat hasil dari metode *K-means Cluster* tersebut UMKM Songkok Nasional Kabupaten Gresik dapat memposisikan perilaku yang sama terhadap penjualan yang ada pada satu klaster yang sama tersebut.

Tabel 2. Analisis Klaster metode *K-means Cluster*

Kota	Klaster 1	Klaster 2
Babat	18,50	65,10
Serang	0,43	3,25
Mojokerto	0,86	7,35
Tumenggung	21,40	17,85
Purwokerto	0,00	50,50

Bojonegoro	18,64	7,40
Tuban	3,57	20,20
Kebumen	38,64	66,50
Cilacap	0,00	104,55
Nganjuk	0,00	37,70
Magelang	68,00	8,00

5. KESIMPULAN

Dari hasil analisis diatas dapat disimpulkan bahwa pemesanan songkok nasional UMKM kabupaten Gresik tidak sama antar konsumen pada kota-kota yang ada. Yang mana hasil analisis klaster tersebut didapat pengelompokan dengan 2 klaster dari sejumlah konsumen pada 11 kota yang ada di Indonesia. Untuk klaster 1 lebih rendah tingkat penjualannya dibandingkan dengan klaster 2.

6. DAFTAR PUSTAKA

- [1] Benri Melpa & Herlina Latipa, ANALISIS CLUSTERING MENGGUNAKAN METODE K-MEANS DALAM PENGELOMPOKAN PENJUALAN PRODUK PADA SWALAYAN FADHILA, Jurnal Media Infotama, Vo. 11. No. 2, 2015
- [2] Johan Oscar, IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI MARKETING PRESIDENT UNIVERSITY, Jurnal Ilmiah Teknik Industri, Vol. 12. No. 1. 2013
- [3] Robinson , Analisis Cluster terhadap Tingkat Pencemaran Udara pada Sektor Industri di Sumatera Selatan, Jurnal Penelitian Sains, Vol 14, No. 3A, 2011
- [4] Safa'at Y & Kishera H ,ANALISIS KLASER UNTUK PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI JAWA TENGAH BERDASARKAN INDIKATOR KESEJAHTERAAN RAKYAT, Jurnal Statistika, Vol 2. No 1, 2014