

# PENERAPAN DATA MINING MENGGUNAKAN METODE CLUSTERING UNTUK EVALUASI EKSPOR IKAN DI INDONESIA MENURUT NEGARA TUJUAN UTAMA

Nazwa Ariana Salsabila, Ultach Enri, Iqbal Maulana

Informatika, Universitas Singaperbangsa Karawang

Jalan Raya Karanglo km 2 Malang, Indonesia

2010631170106@student.unsika.ac.id

## ABSTRAK

Indonesia merupakan salah satu negara yang kaya akan lautan yang biasanya disebut juga sebagai negara maritim. Dimana negara Indonesia yang kaya akan sumber daya laut, hal ini dapat menjadikan negara Indonesia sebagai salah satu negara dengan pengekspor ikan segar terbanyak se-Asia. Namun, pada tahun 2021 nilai ekspor ikan di Indonesia mengalami penurunan dibanding tahun sebelumnya, hal ini diakibatkan ekspor di Indonesia masih belum maksimal meskipun memiliki sumber daya baik itu sumber daya lautnya maupun sumber daya manusia yang banyak. Dimana pada penelitian ini, bertujuan untuk menganalisa dan mengevaluasi ekspor ikan berdasarkan negara tujuan utama dengan menggunakan metode *K-Means Clustering*. Dataset yang akan digunakan yaitu bersumber dari *website* bps. Dimana data tersebut diolah dalam 3 cluster. Hasilnya menunjukkan Hasilnya menunjukkan bahwa ada 2 negara tujuan yang memiliki nilai ekspor terbanyak yaitu negara Malaysia dan Singapura, lalu untuk negara yang memiliki nilai ekspor rata-rata yaitu negara Jepang dan Taiwan, dan sisanya memiliki nilai ekspor rendah.

**Kata kunci :** *Clustering, K-Means, Ekspor Ikan, Data Mining*

## 1. PENDAHULUAN

Indonesia adalah negara yang dikenal dengan negara kepulauan yang memiliki potensi hasil laut yang sangat besar. Potensi ini tersebar di sekitar 5,8 juta km<sup>2</sup> zona maritim yang terdiri dari perairan kepulauan (2,3 juta km<sup>2</sup>), perairan territorial (0,8 juta km<sup>2</sup>), dan Zona Ekonomi Eksklusif (2,7 juta km<sup>2</sup>). Produksi ikan tangkap laut mencapai lebih dari 5 juta ton pada tahun 2012, dengan pertumbuhan sebesar 3,5 persen per tahun selama lebih dari sepuluh tahun (1991 – 2012).[1]

Sebagian produk domestik diperdagangkan di pasar global. Ekspor produk perikanan dan ikan berada di bawah ekspor makanan olahan dan perhiasan, hal ini didapatkan berdasarkan laporan Kementerian Perdagangan. Nilai ekspor produk ini mencapai sekitar 1.278 US\$ pada tahun 2011, dan meningkat menjadi 1.545 US\$ pada tahun 2012. Dengan kata lain, nilai ekspor dalam periode tersebut meningkat sebesar 20,89 persen menurut Kementerian Perdagangan RI pada tahun 2013.

Potensi ekonomi dari hasil laut serta produk terkait masih dapat terus ditingkatkan mengingat luasnya cakupan zona tangkap perikanan laut di Indonesia. Jika di wilayah perairan sebelah barat (termasuk Selat Malaka dan Laut Jawa) mula-mula muncul gejala kelebihan tangkapan (*overfishing*), namun sebaliknya di daerah perairan sebelah timur justru masih sangat kurang intensitas penangkapan ikan [2]. Sumber daya perikanan tangkap memiliki potensi tangkapan lestari maksimum, atau *Maximum Sustainable Yield* (MSY), sekitar 6,4 juta ton per tahun, menurut temuan beberapa penelitian. Dengan

demikian potensi ekspor produk ikan Indonesia masih dapat ditingkatkan di masa mendatang.

Mengingat keunggulan alamiah yang dimiliki oleh Indonesia, maka sangat wajar jika ada harapan bahwa Indonesia seharusnya menjadi pemain utama dalam pasar regional (katakanlah, pada level ASEAN). Meskipun demikian, fakta menunjukkan hal yang berbeda. Saat ini, ekspor komoditas perikanan dan ikan di Indonesia masih kalah dibandingkan dengan ekspor negara tetangga yang memiliki wilayah teritori lebih sempit seperti Singapura, Vietnam dan Thailand. Kondisi ini tentu harus disikapi secara bijak. Berdasarkan latar belakang tersebut, maka penelitian ini akan menganalisis perkembangan negara tujuan ekspor perikanan Indonesia, yang nantinya dapat memberikan masukan kepada pemerintah untuk membantu meningkatkan perekonomian Indonesia khususnya dalam sektor perikanan.

## 2. TINJAUAN PUSTAKA

### 2.1. Penelitian Terdahulu

Penelitian “Penerapan Data Mining pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering” data tersebut diolah menggunakan *Rapidminer* untuk ditentukan nilai *centroid* dalam 3 kelompok (tingkat ekspor tinggi, tingkat ekspor sedang, dan tingkat ekspor rendah). Oleh karena itu, penilaian didasarkan pada indeks ekspor buah-buahan. Indeks ini terdiri dari dua negara yang memiliki *cluster* tingkat ekspor tinggi, yaitu India dan Pakistan; tiga negara yang memiliki *cluster* tingkat ekspor sedang yaitu Singapura, Bangladesh, dan negara lainnya; dan enam negara yang memiliki *cluster* tingkat ekspor rendah yaitu Hongkong,

Tiongkok, Malaysia, Nepal, Vietnam dan Iran. Hasil penelitian ini dapat digunakan untuk menentukan jumlah ekspor buah-buahan menurut negara tujuan.[3]

Penelitian berjudul “Implementasi Data Mining Clustering Data Valuasi Ekspor Kertas Indonesia Menggunakan Algoritma K-Means” memiliki tujuan untuk memperoleh informasi kluster pelabuhan-pelabuhan yang menghasilkan transaksi ekspor kertas yang tinggi. Hasilnya menunjukkan bahwa hanya 10% pelabuhan yang memiliki kategori transaksi ekspor kertas dengan valuasi tinggi, 3% memiliki valuasi sedang, dan 87% atau sekitar 27 pelabuhan, memiliki catatan transaksi level rendah. Selain itu, banyak Pelabuhan ekspor yang berdekatan secara geografis tetapi memiliki nilai transaksi yang sangat berbeda satu sama lain.[4]

Penelitian yang memiliki judul “Penerapan K-Means Clustering pada Ekspor Minyak Kelapa Sawit Menurut Negara Tujuan” memiliki tujuan untuk membagi dua kelompok yaitu kelompok tingkat ekspor tinggi dan kelompok tingkat ekspor rendah. Hasilnya didasarkan pada indeks ekspor minyak kelapa sawit, dengan India sebagai satu-satunya negara kelompok tingkat ekspor tinggi dan Sembilan negara kelompok tingkat ekspor rendah yaitu Tiongkok, Singapura, Malaysia, Pakistan, Bangladesh, Sri Lanka, Mesir, Belanda, dan Jerman. Negara ini memprioritaskan ekspor minyak kelapa sawit cluster.[5]

## 2.2. Clustering

Clustering dikenal sebagai teknik yang dapat membagi objek dari kumpulan banyak data menjadi beberapa kelompok yang sama. Teknik pengelompokan pada dasarnya adalah mengelompokkan berbagai artikel atau informasi ke dalam kelompok atau kumpulan yang berbeda, sehingga setiap kelompok mengandung informasi yang diharapkan dapat dibandingkan (komparatif). Clustering digunakan untuk menyatukan data yang serupa dalam satu kelompok dan mengurangi jarak antar kelompok. Ini menunjukkan bahwa barang dalam setiap kelompok tidak sama dengan barang dalam kelompok lain, tetapi sangat mirip satu sama lain.[6]

## 2.3. K-Means Clustering

Metode *K-Means* yaitu metode yang membagi informasi ke dalam kelompok-kelompok sehingga kumpulan yang memiliki kualitas yang sebanding dan kumpulan yang memiliki kualitas yang berbeda dikumpulkan ke dalam kumpulan yang berbeda. Berikut langkah-langkah perhitungannya : [7]

- Tentukan jumlah klaster yang diinginkan dengan cara menghitung nilai  $k$ .
- Menginisialisasi  $k$  pusat dengan menggunakan teknik acak yang diperoleh dari data sebelumnya.
- Cari data yang paling dekat dengan *centroid* dengan menggunakan persamaan jarak Euclidean.

Berikut adalah persamaan *Euclidian Distance* ditunjukkan pada persamaan (1) :

$$D(x_i, x_j) = \sqrt{(x_{i1} - x_{j1})^2 \dots + (x_{ki} - x_{kj})^2} \quad (1)$$

Dimana :

$D(x_i, x_j)$  = Jarak data  $i$  ke pusat *cluster*  $j$

$X_{ki}$  = Data  $i$  pada pada atribut data ke  $k$

$X_{kj}$  = Titik pusat  $j$  pada atribut ke  $k$

- Urutkan setiap data berdasarkan jarak terdekatnya dengan *centroid* (jarak terpendek).
- Perbarui nilai *centroid*; nilai *centroid* terbaru diperoleh dari *cluster* rata-rata dengan menggunakan rumus yang ditunjukkan pada persamaan (2) di bawah ini :

$$C_{m(q)} = \frac{1}{n_m} \sum_{i=1}^{n_m} x_{i(q)} \quad (2)$$

Dimana :

$C_m(q)$  = Centroid dari klaster  $m$  pada iterasi ke- $q$

$N_m$  = Jumlah data yang termasuk dalam klaster  $m$

$X_i(q)$  = Data ke- $i$  pada iterasi ke- $q$

- Ulangi tahap 2 hingga 5 sampai semua individu tidak berubah.

## 2.4. Data Mining

Setiap harinya manusia memproduksi data, maka dari itu ada banyak data di dunia pada saat ini. Data yang diperoleh sangat beragam macamnya, termasuk data pemerintah, data keuangan, data Pendidikan, dan lain-lain yang berasal dari berbagai industri. Data yang beragam ini diolah untuk menghasilkan pengetahuan baru. *Data mining* proses analisis data untuk menemukan hubungan yang jelas dan menghasilkan kesimpulan yang belum diketahui sebelumnya dengan cara yang mudah dipahami dan bermanfaat bagi siapa saja yang memiliki data. Saat ini data mining dibagi menjadi beberapa bagian diantaranya seperti *K-Means*, *Naive Bayes*, *C4.5*, *NVM* (*Support Vector Machine*).[8]

## 3. METODE PENELITIAN

Untuk mendapatkan hasil penelitian yang optimal, pemilihan metode penelitian sangat penting dalam sebuah riset. Penelitian yang akan dilakukan ini menggunakan Algoritma *K-Means Clustering* untuk melakukan evaluasi ekspor ikan di Indonesia menurut negara tujuan utama.

Tujuan dari *clustering* merupakan sebuah teknik penganalisa data yang banyak digunakan sebagai salah satu teknik *data mining* yang bertujuan untuk mengelompokkan data dengan atribut yang sebanding. Salah satu metode *data mining* yang dikenal dengan *clustering* ini berfungsi untuk meletakkan data pada kelompok atau *cluster* tertentu [9]. Setiap data dikelompokkan berdasarkan seberapa banyak kemiripannya dengan data yang lain. *Cluster* yang terdiri dari data dengan atribut yang sama akan digabungkan, dan *cluster* yang terdiri dari data dengan atribut yang berbeda akan digabungkan. *K-Means* adalah salah satu dari banyaknya metode *clustering* yang tersedia.

*K-Means* adalah salah satu metode atau pendekatan *clustering* yang paling umum diaplikasikan. Algoritma *K-Means* adalah salah satu metode *clustering* non-hirarki dengan waktu komputasi yang cukup cepat dibanding dengan metode *clustering* yang lain [10]. Metode algoritma *K-Means* ini awalnya dimulai dengan menentukan berapa banyak jumlah *cluster* yang akan dibuat, selanjutnya hitung nilai awal setiap *cluster*, dan kemudian menghitung jarak yang ada diantara masing-masing data dengan nilai awal tersebut. Setelah menghitung jarak masing-masing data, data dimasukkan ke dalam *cluster* terdekat. Selanjutnya, kita menghitung rata-rata dari setiap *cluster* tersebut, yang kemudian akan digunakan sebagai nilai awal yang baru untuk menghitung jarak dari masing-masing *cluster*.

Pada metode *K-Means* ini, *centroid* atau pusat *cluster* dipilih secara acak dari sekumpulan koleksi atau populasi informasi pada tahap awal. Setelah itu, *K-Means* memeriksa setiap komponen dalam populasi data dan menandai setiap komponen ke salah satu *centroid* yang telah ditetapkan sebelumnya, berdasarkan jarak minimum antara komponen (data) dan *centroid*. Sampai semua komponen atau data telah terkumpul semua di setiap *centroid*, posisi *centroid* akan dihitung kembali dan terakhir sampai semua komponen (data).

Untuk melakukan *clustering* dengan metode *K-Means Clustering*, berikut langkah-langkah yang harus diikuti:

1. Menghitung nilai K untuk jumlah *cluster* yang diinginkan untuk dibentuk.
2. Memilih *centroid* atau pusat *cluster* awal dengan nilai K..
3. Dengan menggunakan rumus jarak *Euclidean*, akan ditemukannya jarak yang paling dekat dari setiap data ke *centroid*.
- Jarak *Euclidean* yang dirumuskan sebagai berikut
 
$$D(i,j) = \sqrt{(X1j - X1i)^2 + (X2j - X2i)^2 + \dots + (Xkj - Xki)^2} \dots \quad (3)$$

D(i,j) = Jarak data ke i ke pusat *cluster* j  
 Xki = Data ke i pada atribut data ke K  
 Xkj = Data ke j pada atribut data ke K

#### 4. HASIL DAN PEMBAHASAN

Data yang didapatkan didapatkan dari website <https://www.bps.go.id>. Dimana data tersebut yang akan dijadikan sampel dan di uji cobakan.

##### 4.1. Data Preparing

Sebelum melakukan pengujian, kami mengambil data sampel dari *website* resmi BPS. Dimana data tersebut berisikan data ekspor ikan berdasarkan negara tujuan pada tahun 2012-2021.

Negara tujuan	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
0 Jepang	12716.3	11921.1	11050.0	6191.5	6166.9	4335.2	2862.8	2265.4	3025.9	2781.4
1 Malaysia	26174.2	27756.2	25538.0	30721.2	25110.9	23384.7	26619.4	28359.8	30225.9	30861.9
2 Singapura	20014.5	24181.6	22081.6	22924.4	19760.3	18143.6	17528.1	15452.2	15591.1	15235.2
3 Amerika Serikat	640.0	316.8	383.2	641.6	701.1	287.8	241.5	171.0	101.6	22.9
4 Taiwan	3844.6	3057.7	2721.2	2669.8	2361.7	2041.0	1906.9	2311.3	2448.9	1740.6
5 Arab Saudi	1.1	58.2	99.0	503.2	1131.9	886.3	809.4	697.8	262.1	105.9
6 Hongkong	936.1	1002.5	1298.6	1425.3	1917.4	2332.2	2341.2	2124.3	2519.9	1074.6
7 Tiongkok	13919.7	8740.4	3202.6	461.0	1121.3	897.9	1725.7	1874.1	2817.4	3151.5
8 Australia	571.2	710.8	457.4	552.3	432.5	378.8	367.4	320.8	271.7	197.0
9 Uni Emirat Arab	139.6	54.6	47.9	150.3	462.0	722.6	298.9	144.5	58.6	7.7
10 Lainnya	12743.8	20279.1	1099.3	977.9	818.0	571.1	260.1	185.5	571.2	138.0

Gambar 1. Data ekspor Ikan Indonesia Tahun 2012-2021

##### 4.2. Data Understanding

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11 entries, 0 to 10
Data columns (total 11 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   Negara tujuan   11 non-null     object
1   2012            11 non-null     float64
2   2013            11 non-null     float64
3   2014            11 non-null     float64
4   2015            11 non-null     float64
5   2016            11 non-null     float64
6   2017            11 non-null     float64
7   2018            11 non-null     float64
8   2019            11 non-null     float64
9   2020            11 non-null     float64
10  2021            11 non-null     float64
dtypes: float64(10), object(1)
memory usage: 1.1+ KB
```

Gambar 1. Hasil dari fungsi info() pada pandas

Pada tahap ini, kita harus memastikan terlebih dahulu bahwa data yang didapat dan dipakai untuk uji coba harus sesuai dengan kualifikasi. Untuk melakukannya, kami menggunakan *library Pandas*. Dengan menggunakan fungsi `info()` untuk mengecek tipe data, nama kolom, dan mengecek apakah ada data yang kosong.

Dapat dilihat pada gambar, dimana data yang didapatkan memiliki 11 kolom. Kolom “Negara tujuan” memiliki tipe data *object* dan pada kolom, 2012-2021 memiliki tipe data *float*. Selain itu, pada setiap kolom juga tidak memiliki data kosong atau null.

Untuk lebih jelasnya, data tersebut ditampilkan dalam bentuk grafik. Dengan menggunakan *library matplotlib*, dapat menampilkan data ke dalam bentuk grafik dimana kodenya sebagai berikut.

```
from matplotlib import pyplot as plt
from matplotlib.pyplot import figure

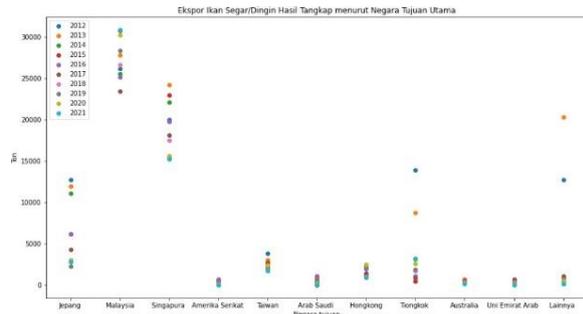
arr = []
figure(figsize=(15, 8))
for i in range(1, 11):
    plt.scatter(df['Negara tujuan'], df[str(2011 + i)])
    arr.append(str(2011 + i))

plt.title('Ekspor Ikan Segar/Dingin Hasil Tangkap menurut Negara Tujuan Utama')
plt.xlabel('Negara tujuan')
plt.ylabel('Ton')

plt.legend(arr)
```

Gambar 2. Kode untuk menampilkan grafik

Dan untuk hasilnya dapat dilihat pada gambar di bawah ini.



Gambar 3. Nilai ekspor ikan Indonesia tahun 2012-2021

Dapat dilihat bahwa negara Malaysia merupakan negara penerima ekspor dari Indonesia ke satu dan bahkan nilai ekspor nya selalu naik setiap tahunnya. Namun, untuk beberapa negara seperti Arab Saudi, Hongkong, dan Tiongkok nilai ekspor nya menurun dari tahun sebelumnya ke tahun 2021.

4.3. Data Transformation

Dimana sebelum melakukan uji coba, akan dilakukan penghitungan rata-rata nilai ekspor ikan dari setiap negara tujuan dengan jangka tahun 2012-2021. Dimana untuk menghitung rata-ratanya, digunakan kode sebagai berikut.

Dan untuk hasilnya dapat dilihat pada gambar di bawah ini.

```
# get average value
average = []

copy = df.copy()

for i in range(0, 11):
    count = 0
    for j in range(1, 11):
        count += df[str(2011 + j)][i]

    average.append(count/10)

copy['average'] = average

copy
```

Gambar 4. Kode untuk mendapatkan nilai rata-rata ekspor ikan tahun 2012-2021

Dikarenakan data yang didapat dari website BPS pada value ekspor masih dalam bentuk ton, disini kami men-scale data tersebut menggunakan MinMaxScaler. hal ini dilakukan dikarenakan angka nya masih besar dan hal ini dapat berpengaruh pada model yang akan dibuat. Dimana kami melakukannya dengan menggunakan rumus berikut.

Negara tujuan	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	average
0 Jepang	12716.3	11921.1	11050.0	6191.5	6166.9	4335.2	2862.8	2265.4	3025.9	2781.4	6331.65
1 Malaysia	26174.2	27756.2	25538.0	30721.2	25110.9	23384.7	26619.4	28359.8	30225.9	30861.9	27475.22
2 Singapura	20014.5	24161.6	22081.6	22924.4	19760.3	18143.6	17528.1	16452.2	15591.1	15232.5	19091.26
3 Amerika Serikat	640.0	318.8	383.2	641.6	701.1	287.8	241.5	171.0	101.6	22.9	350.75
4 Taiwan	3844.6	3057.7	2721.2	2669.8	2361.7	2041.0	1906.9	2311.3	2448.9	1740.6	2510.37
5 Arab Saudi	1.1	58.2	99.0	503.2	1131.9	886.3	809.4	697.8	262.1	105.9	455.49
6 Hongkong	936.1	1002.5	1298.6	1425.3	1917.4	2333.2	2341.2	2124.3	2519.9	1074.6	1697.31
7 Tiongkok	13919.7	8740.4	3202.6	461.0	1121.3	897.9	1725.7	1874.1	2617.4	3151.5	3771.16
8 Australia	571.2	710.8	457.4	552.3	432.5	378.8	367.4	320.8	271.7	197.0	425.99
9 Uni Emirat Arab	139.6	54.6	47.9	150.3	462.0	722.6	298.9	144.5	58.6	7.7	208.67
10 Lainnya	12743.8	20279.1	1099.3	977.9	818.0	571.1	260.1	185.5	571.2	138.0	3764.40

Gambar 5. Rata-rata ekspor ikan Indonesia tahun 2012-2021

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Dimana untuk melakukannya, digunakan sklearn untuk mempermudah proses, kode nya dapat dilihat dibawah ini.

```
from sklearn.preprocessing import MinMaxScaler

arr.append('average')

scaler = MinMaxScaler()
scale = scaler.fit_transform(copy[arr])
df_scale = pd.DataFrame(scale, columns = arr);

df_scale.insert(0, 'Negara Tujuan', df['Negara tujuan'])

df_scale
```

Gambar 6. Kode untuk scale nilai ekspor

Setelah di scale menjadi angka antara 0-1 dengan menggunakan MinMaxScaler, hasilnya dapat dilihat pada tabel di bawah ini.

Negara Tujuan	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	average
0 Jepang	0.485812	0.428369	0.431622	0.197613	0.232365	0.175236	0.099375	0.075168	0.098361	0.089897	0.224560
1 Malaysia	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
2 Singapura	0.764655	0.870961	0.864402	0.744960	0.783187	0.773082	0.655344	0.642532	0.514879	0.493531	0.692518
3 Amerika Serikat	0.024411	0.009465	0.013154	0.016071	0.010884	0.000000	0.000000	0.000939	0.001425	0.000493	0.005211
4 Taiwan	0.146849	0.108409	0.104878	0.082415	0.078174	0.075906	0.063136	0.076795	0.079235	0.056164	0.084415
5 Arab Saudi	0.000000	0.000130	0.002005	0.011544	0.026341	0.025913	0.021529	0.019610	0.006748	0.003183	0.009052
6 Hongkong	0.035724	0.034218	0.049066	0.041706	0.060170	0.088567	0.079601	0.070168	0.081588	0.034579	0.054506
7 Tiongkok	0.531790	0.313549	0.123762	0.010163	0.027911	0.026415	0.056267	0.061300	0.084820	0.101892	0.130654
8 Australia	0.021782	0.023688	0.016065	0.010150	0.000000	0.003940	0.004773	0.006248	0.007064	0.006135	0.007970
9 Uni Emirat Arab	0.005292	0.000000	0.000000	0.000000	0.001195	0.018825	0.002176	0.000000	0.000000	0.000000	0.000000
10 Lainnya	0.486862	0.730084	0.041247	0.027071	0.015821	0.012266	0.000705	0.001453	0.016992	0.004223	0.130406

Gambar 7. Nilai ekspor setelah di scale

4.4. Membuat Model

Pada tahap ini yaitu membuat model clustering menggunakan dataset yang telah ditransformasi. Untuk melakukannya, digunakan library sklearn. Dengan cara membagnya menjadi 3 cluster dengan random state 1. Untuk hasilnya dapat dilihat pada gambar di bawah.

Negara Tujuan	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	average	clusters
0 Jepang	0.485812	0.428369	0.431622	0.197613	0.232365	0.175236	0.099375	0.075168	0.098361	0.089897	0.224560	2
1 Malaysia	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1
2 Singapura	0.764655	0.870961	0.864402	0.744960	0.783187	0.773082	0.655344	0.642532	0.514879	0.493531	0.692518	1
3 Amerika Serikat	0.024411	0.009465	0.013154	0.016071	0.010884	0.000000	0.000000	0.000939	0.001425	0.000493	0.005211	0
4 Taiwan	0.146849	0.108409	0.104878	0.082415	0.078174	0.075906	0.063136	0.076795	0.079235	0.056164	0.084415	0
5 Arab Saudi	0.000000	0.000130	0.002005	0.011544	0.026341	0.025913	0.021529	0.019610	0.006748	0.003183	0.009052	0
6 Hongkong	0.035724	0.034218	0.049066	0.041706	0.060170	0.088567	0.079601	0.070168	0.081588	0.034579	0.054506	0
7 Tiongkok	0.531790	0.313549	0.123762	0.010163	0.027911	0.026415	0.056267	0.061300	0.084820	0.101892	0.130654	2
8 Australia	0.021782	0.023688	0.016065	0.010150	0.000000	0.003940	0.004773	0.006248	0.007064	0.006135	0.007970	0
9 Uni Emirat Arab	0.005292	0.000000	0.000000	0.000000	0.001195	0.018825	0.002176	0.000000	0.000000	0.000000	0.000000	0
10 Lainnya	0.486862	0.730084	0.041247	0.027071	0.015821	0.012266	0.000705	0.001453	0.016992	0.004223	0.130406	2

Gambar 8. Cluster nilai ekspor ikan Indonesia tahun 2012-2021

Dikarenakan variabel yang digunakan pada saat membuat model lebih dari 2, diperlukan mereduksi dimensi data kita. Dimana kita bisa menggunakan *library sklearn* untuk melakukannya dengan kode dibawah.

```
from sklearn.decomposition import PCA

pca = PCA(n_components=2).fit_transform(df_scale[arr])
pcadf = pd.DataFrame({
    'X': pca[:, 0],
    'Y': pca[:, 1],
    'clusters': df_scale['clusters']
})

pcadf
```

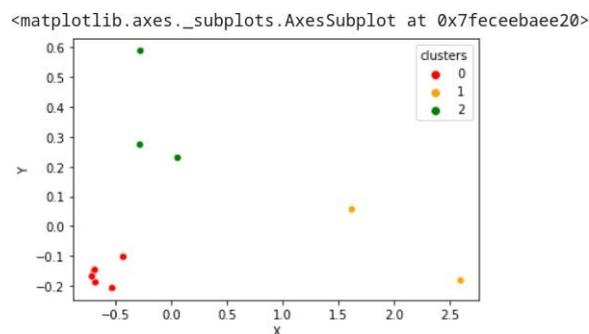
Gambar 9. Kode untuk melakukan PCA

Dan hasilnya dapat dilihat pada gambar di bawah.

	Negara Tujuan	X	Y	clusters
0	Jepang	0.058708	0.229111	2
1	Malaysia	2.594939	-0.182087	1
2	Singapura	1.620110	0.056394	1
3	Amerika Serikat	-0.692179	-0.151772	0
4	Taiwan	-0.428642	-0.102922	0
5	Arab Saudi	-0.678404	-0.187903	0
6	Hongkong	-0.527828	-0.207203	0
7	Tiongkok	-0.279551	0.273272	2
8	Australia	-0.683834	-0.146419	0
9	Uni Emirat Arab	-0.708862	-0.168249	0
10	Lainnya	-0.274456	0.587779	2

Gambar 10. Hasil clustering setelah PCA

Untuk hasil yang lebih jelas, kita dapat melihat grafik di bawah ini.



Gambar 11. Grafik nilai ekspor ikan Indonesia tahun 2012-2021

Dapat dilihat bahwa ada 2 negara tujuan yang memiliki nilai ekspor ikan terbanyak yaitu negara Malaysia dan Singapura, lalu untuk negara yang

memiliki nilai ekspor rata-rata yaitu Jepang dan Taiwan. dan sisanya memiliki nilai ekspor rendah.

### 5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dengan menggunakan data yang dalam penelitian ini pada tahun 2012-2021, dapat diambil kesimpulan bahwa hasil penelitian ini menunjukkan bahwa metode pengelompokan *K-Means* dapat digunakan untuk menentukan pengelompokan ekspor ikan. Data ini telah diolah dengan mengelompokkan dengan sebanyak dua kelompok, yaitu kelompok tingkat ekspor tinggi dan kelompok tingkat ekspor rendah. Hasilnya menunjukkan bahwa ada 2 negara tujuan yang memiliki nilai ekspor terbanyak yaitu negara Malaysia dan Singapura, lalu untuk negara yang memiliki nilai ekspor rata-rata yaitu negara Jepang dan Taiwan, dan sisanya memiliki nilai ekspor rendah.

Pengolahan data untuk *clustering* memiliki kemampuan untuk memberikan pembobotan kriteria untuk membuat data yang dihasilkan lebih akurat. Selain itu, pengolahan data untuk *clustering* juga memerlukan peningkatan akurasi data yang dikumpulkan. Untuk mencapai hasil yang lebih akurat dari proses *clustering*, kriteria yang lebih spesifik dapat ditambahkan.

### DAFTAR PUSTAKA

- [1] K. B. Syariah and G. Ilmu, "Perkembangan Beberapa Indikator Utama Sosial-Ekonomi Indonesia".
- [2] Y. I. Siregar, "Menggali Potensi Sumberdaya Laut Indonesia," *Makalah*, 2015.
- [3] A. P. Windarto, "Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Method," *Techno.Com*, vol. 16, no. 4, pp. 348–357, 2017, doi: 10.33633/tc.v16i4.1447.
- [4] N. Pooja, M. Saputra, S. Aisyah, and P. Juanta, "Implementasi Data Mining Clustering Data Valuasi Ekspor Kertas Indonesia Menggunakan Algoritma K-Means," *J. Sist. Inf. dan Ilmu Komputer. Prima (JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 86–90, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2372.
- [5] S. Hajar, A. A. Novany, A. P. Windarto, A. Wanto, and E. Irawan, "Penerapan K-Means Clustering pada ekspor minyak kelapa sawit menurut negara tujuan," *Semin. Nas. Teknol. Komput. Sains 2020*, pp. 314–318, 2020.
- [6] A. Bahauddin, A. Fatmawati, and F. Permata Sari, "Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means," *J. Manaj. Inform. dan Sist. Inf.*, vol. 4, no. 1, pp. 1–8, 2021, doi: 10.36595/misi.v4i1.216.
- [7] A. N. Alifia, A. F. Setiawan, and D. Rudhistiar, "PENERAPAN ALGORITMA K-MEANS CLUSTERING DALAM PERINGATAN DINI

- RESIKO KEGAGALAN SISWA PADA MATA PELAJARAN BAHASA INDONESIA ( STUDI KASUS : SMP NEGERI 2 MALANG ),” vol. 8, no. 2, pp. 1174–1181, 2024.
- [8] M. Anjelita, A. P. Windarto, A. Wanto, and S. Saifullah, “Analisis Metode K-Means pada Kasus Ekspor Barang Perhiasan dan Barang Berharga Berdasarkan Negara Tujuan,” *Semin. Nas. Sains Teknol. Inf.*, pp. 476–482, 2019, [Online]. Available: <http://prosiding.seminar-id.com/index.php/sensasi/issue/archivePage%7C476>
- [9] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [10] T. Amalina, D. Bima, A. Pramana, and B. N. Sari, “Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food,” *J. Ilm. Wahana Pendidik.*, vol. 8, no. 15, pp. 574–583, 2022, [Online]. Available: <https://doi.org/10.5281/zenodo.7052276>