

# PENGGEROMBOLAN SUBSEKTOR INDUSTRI BERDASARKAN PERKEMBANGAN INDEKS PRODUKSI MENGGUNAKAN PREDICTION-BASED CLUSTERING\*

Agustin Faradila<sup>1</sup>, Utami Dyah Syafitri<sup>2‡</sup>, I Made Sumertajaya<sup>3</sup>

<sup>1</sup>Badan Pusat Statistik, Indonesia, faradila\_agustin@apps.ipb.ac.id

<sup>2</sup>Department of Statistics, IPB University, Indonesia, utamids@apps.ipb.ac.id

<sup>3</sup>Department of Statistics, IPB University, Indonesia, imsjaya.stk@gmail.com

‡corresponding author

**Indonesian Journal of Statistics and Its Applications (eISSN:2599-0802)**

**Vol 4 No 3 (2020), 419 - 431**

Copyright © 2020 Agustin Faradila, Utami Dyah Syafitri, I Made Sumertajaya. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## **Abstract**

Statistics Indonesia (BPS) noted that there has been a decrease in the contribution of the industrial sector to the national GDP even though it had provided a significant multiplier effect on national economic growth. Therefore, it is necessary to cluster the industrial subsector based on its growth patterns so that the optimization of development results can be achieved. Prediction-based clustering is part of time series clustering (TSclust) which aims to form clusters based on prediction characteristics so that it can be used to choose a cluster that will become a mainstay industry in the future. This study focused on applying prediction-based clustering in the large and medium industrial subsector for a prediction period of 1 month, 1 quarter, and 1 semester. The data used in this study was the production index data from January 2010 to December 2018. The results showed that the best cluster for 1 month consisted of 5 groups, for 1 quarter consisted of 4 groups and for 1 semester consisted of 2 groups. Thus, it was concluded that the food industry; leather industry, leather goods, and footwear; and the pharmaceutical industry, chemical drug products, and traditional medicine could be chosen to be the mainstay industry in the future.

**Keywords:** clustering, industry, prediction-based, TSclust.

---

\* Received Nov 2019; Accepted Oct 2020; Published online on Nov 2020

## 1. Pendahuluan

Data BPS menunjukkan bahwa sepanjang tahun 2010 hingga 2018 kontribusi industri pengolahan terhadap PDB nasional menunjukkan tren menurun, padahal sektor industri memberikan multiplier effect yang cukup besar bagi pertumbuhan ekonomi nasional, yaitu penerimaan devisa dari ekspor dan pajak serta penyerapan tenaga kerja yang cukup banyak. Pada saat yang bersamaan, Kementerian Perindustrian (Kemenperin) menargetkan rata-rata pertumbuhan yang harus dicapai oleh sektor industri pada tahun 2010-2020 adalah sebesar 6.34% sementara anggaran pembangunan yang diperlukan sektor industri terbatas. Dengan demikian, perlu dipilih industri yang akan menjadi industri andalan masa depan agar optimalisasi hasil pembangunan dapat tercapai. Salah satu upayanya adalah dengan membuat penggerombolan subsektor industri berdasarkan pola pertumbuhannya. Indikator utama untuk mengetahui pertumbuhan industri tersebut adalah indeks produksi yang diukur secara konsisten dari waktu ke waktu.

*Clustering* (penggerombolan) bertujuan mengelompokkan objek individu berdasarkan pada kemiripan atau ketidakmiripan dari objek-objek yang diteliti menggunakan konsep jarak. Sementara itu, *time series clustering* (penggerombolan deret waktu) merupakan bagian dari *clustering* dengan data berupa data deret waktu. Isu utama dalam penggerombolan deret waktu adalah konsep ketidakmiripan yang menjadi kompleks karena karakteristik data deret waktu yang dinamis. Liao (2005) membagi penggerombolan deret waktu menjadi tiga pendekatan utama, yaitu *raw-based clustering* yang didasarkan secara langsung pada *raw data* (data asli); *feature-based clustering* yang didasarkan secara tidak langsung pada representasi data yang diekstraksi dari *raw data* seperti misalnya dengan *autocorrelation function* (ACF); dan *model-based clustering* yang didasarkan secara tidak langsung pada model yang dibangun dari *raw data* seperti ARIMA.

Ketiga pendekatan penggerombolan tersebut didasarkan oleh penggerombolan pada periode waktu observasi. Dalam beberapa situasi praktis, kepentingan sebenarnya dari penggerombolan bergantung pada sifat prediksi. Alonso et al. (2006) menginisiasi *prediction-based clustering* yang berguna untuk melihat konvergensi/divergensi jangka panjang dan untuk mengetahui apakah level tertentu akan tercapai di masa mendatang. Metode ini menggunakan jarak berdasarkan fungsi kepekatan prediksi yang diperoleh dari prosedur *sieve bootstrap*. Prosedur tersebut memerlukan asumsi data deret waktu mengikuti fungsi parametrik *autoregressive* (AR). Vilar et al. (2010) mengembangkan *prediction-based clustering* untuk model nonparametrik dengan menggunakan prosedur *autoregression bootstrap* tanpa mengasumsikan model parametrik untuk struktur AR yang sebenarnya.

Menentukan ukuran ketidakmiripan yang tepat harus disesuaikan dengan tujuan dari penggerombolan (Montero & Vilar, 2014). Tujuan penggerombolan pada penelitian ini adalah memilih industri andalan yang dapat memenuhi target pertumbuhan di masa mendatang, sehingga pendekatan *clustering* yang sesuai adalah *prediction-based clustering*. Oleh karena itu tujuan penelitian ini adalah menerapkan penggerombolan subsektor industri menggunakan *prediction-based* berdasarkan perkembangan indeks produksi. Penelitian ini dibatasi hanya pada industri berskala besar dan sedang (IBS).

## 2. Metodologi

### 2.1 Bahan dan Data

Data yang digunakan dalam penelitian ini adalah data sekunder hasil publikasi BPS, yaitu Perkembangan Indeks Produksi Industri Manufaktur. Data indeks produksi yang digunakan dalam penelitian ini menggunakan tahun dasar 2010. Data yang dikumpulkan merupakan data bulanan dari Januari 2010 sampai Desember 2018 untuk 23 subsektor menurut KBLI 2 digit disajikan pada Tabel 1.

Tabel 1: Subsektor IBS menurut KBLI 2 *digit*

Peubah	Kode KBLI	Uraian
X <sub>10</sub>	10	Industri Makanan
X <sub>11</sub>	11	Industri Minuman
X <sub>12</sub>	12	Industri Pengolahan Tembakau
X <sub>13</sub>	13	Industri Tekstil
X <sub>14</sub>	14	Industri Pakaian Jadi
X <sub>15</sub>	15	Industri Kulit, Barang dari Kulit dan Alas Kaki
X <sub>16</sub>	16	Industri Kayu, Barang dari Kayu dan Gabus (Tidak Termasuk Furnitur) dan Barang Anyaman dari Bambu, Rotan dan Sejenisnya
X <sub>17</sub>	17	Industri Kertas dan Barang dari Kertas
X <sub>18</sub>	18	Industri Pencetakan dan Reproduksi Media Rekaman
X <sub>20</sub>	20	Industri Bahan Kimia dan Barang dari Bahan Kimia
X <sub>21</sub>	21	Industri Farmasi, Produk Obat Kimia dan Obat Tradisional
X <sub>22</sub>	22	Industri Karet, Barang dari Karet dan Plastik
X <sub>23</sub>	23	Industri Barang Galian Bukan Logam
X <sub>24</sub>	24	Industri Logam Dasar
X <sub>25</sub>	25	Industri Barang Logam, Bukan Mesin dan Peralatannya
X <sub>26</sub>	26	Industri Komputer, Barang Elektronik dan Optik
X <sub>27</sub>	27	Industri Peralatan Listrik
X <sub>28</sub>	28	Industri Mesin dan Perlengkapan ytdl
X <sub>29</sub>	29	Industri Kendaraan Bermotor, Trailer dan Semi Trailer
X <sub>30</sub>	30	Industri Alat Angkutan Lainnya
X <sub>31</sub>	31	Industri Furnitur
X <sub>32</sub>	32	Industri Pengolahan Lainnya
X <sub>33</sub>	33	Jasa Reparasi dan Pemasangan Mesin dan Peralatan

### 2.2 Metode Penelitian

Analisis data dilakukan menggunakan *software R studio* dengan *package TSclust*. Berikut tahapan-tahapan yang dilakukan dalam penelitian ini:

1. Melakukan eksplorasi data untuk mengetahui gambaran mengenai pola perkembangan indeks produksi IBS pada masing-masing subsektor.
2. Memeriksa kestasioneran data indeks produksi pada 23 subsektor dengan menggunakan uji *Augmented Dickey-Fuller* (ADF) dan uji *Levene*.
3. Melakukan prosedur *autoregression bootstrap* untuk mendapatkan dugaan fungsi

kepekatan prediksi pada setiap *series* subsektor ke- $i$ ,  $X_t^{(i)}$ , dengan tahapan sebagai berikut (Vilar et al., 2010):

- 1) Untuk setiap *series*, suatu proses stasioner  $X_t$  berlaku model

$$X_t = m(X_{t-1}) + \varepsilon_t \quad (1)$$

Langkah pertama adalah menduga fungsi  $m$  menggunakan penduga kernel. Metode kernel merupakan salah satu teknik nonparametrik yang bertujuan untuk menduga fungsi kepekatan dari data pengamatan, dan merupakan pengembangan dari histogram. Penduga kernel diperkenalkan oleh Rosenblatt tahun 1956 dan Parzen tahun 1962 sehingga disebut penduga kepekatan kernel Rosenblatt-Parzen (Hardle, 1994). Penduga kernel Rosenblatt-Parzen dengan *bandwidth*  $h$  didefinisikan sebagai berikut:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) \quad (2)$$

dimana

$$K_h(u) = \frac{1}{h} K\left(\frac{u}{h}\right) \text{ untuk } -\infty < x < \infty, h > 0 \quad (3)$$

Ada beberapa jenis fungsi kernel, namun pemilihan *bandwidth* jauh lebih penting dibandingkan memilih fungsi kernel karena penggunaan fungsi kernel yang berbeda dengan nilai *bandwidth* optimal yang sama menghasilkan dugaan kurva yang hampir sama (Sukarsa & Srinadi, 2012). Fungsi kernel yang akan digunakan dalam penelitian ini adalah *Gaussian kernel* karena yang paling umum dan mudah diaplikasikan. Fungsi *Gaussian kernel* didefinisikan sebagai berikut:

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) \quad (4)$$

Dalam pendekatan nonparametrik, bentuk kurva hubungan antara dua peubah tidak diketahui. Regresi kernel digunakan untuk menduga fungsi regresi  $m$  pada model regresi nonparametrik deret waktu dalam penelitian ini  $X_t = m(X_{t-1}) + \varepsilon_t$ . Nadaraya dan Watson pada tahun 1964 mendefinisikan penduga regresi kernel sehingga disebut penduga Nadaraya-Watson, yaitu:

$$\hat{m}_h(x) = \frac{n^{-1} \sum_{t=2}^n K_h(x - X_{t-1}) X_t}{n^{-1} \sum_{t=2}^n K_h(x - X_t)} \quad (5)$$

Pada langkah pertama ini,  $m$  pada persamaan (1) diduga menggunakan persamaan (5) dengan *bandwidth*  $h = g_1$  yang dipilih menggunakan *time series cross validation*

- 2) Menghitung sisaan  $\hat{\varepsilon}_t = X_t - \hat{m}_{g_1}(X_{t-1})$
- 3) Membangun fungsi  $\hat{f}_{\hat{\varepsilon}, h}$  menggunakan penduga Rosenblatt-Parzen  $\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i)$  dengan *bandwidth*  $h$  dipilih menggunakan *plug-in selector* oleh Sheather & Jones (1991)
- 4) Menarik *bootstrap-resample*  $\varepsilon_t^*$  dari  $\hat{f}_{\hat{\varepsilon}, h}$
- 5) Mendefinisikan *series*  $X_t^* = \hat{m}_{g_1}(X_{t-1}^*) + \varepsilon_t^*$ ;  $t = 1, \dots, T$
- 6) Memperoleh  $\hat{m}_{g_2}^*$  dari sampel *bootstrap*  $(X_1^*, \dots, X_T^*)$  menggunakan penduga Nadaraya-Watson dengan *bandwidth*  $g_2 = 1.5g_1$  atau  $g_2 = 2g_1$
- 7) Memperoleh  $X_t^* = \hat{m}_{g_2}^*(X_{t-1}^*) + \varepsilon_t^*$  untuk  $t = T + 1, \dots, T + b$
- 8) Mengulangi langkah 1) sampai 7) sebanyak  $B=1000$  kali untuk mendapatkan replikasi dari observasi *bootstrap* pada  $b$  waktu mendatang. Pada penelitian ini digunakan  $b=1$  untuk jangka waktu prediksi 1 bulan,  $b=3$  untuk jangka

waktu prediksi 3 bulan (1 triwulan), dan  $b=6$  untuk jangka waktu prediksi 6 bulan (1 semester).

4. Membangun dugaan fungsi kepekatan prediksi  $\hat{f}_{X_{t+b}}(u)$  menggunakan penduga Rosenblatt-Parzen dengan *bandwidth*  $h$  dipilih menggunakan *plug-in selector* oleh Sheather & Jones (1991)

5. Menghitung matriks ketidakmiripan dengan jarak prediksi

$$d_{PRED,b}(X_t, Y_t) = \int |\hat{f}_{X_{t+b}}(u) - \hat{f}_{Y_{t+b}}(u)| du \quad (6)$$

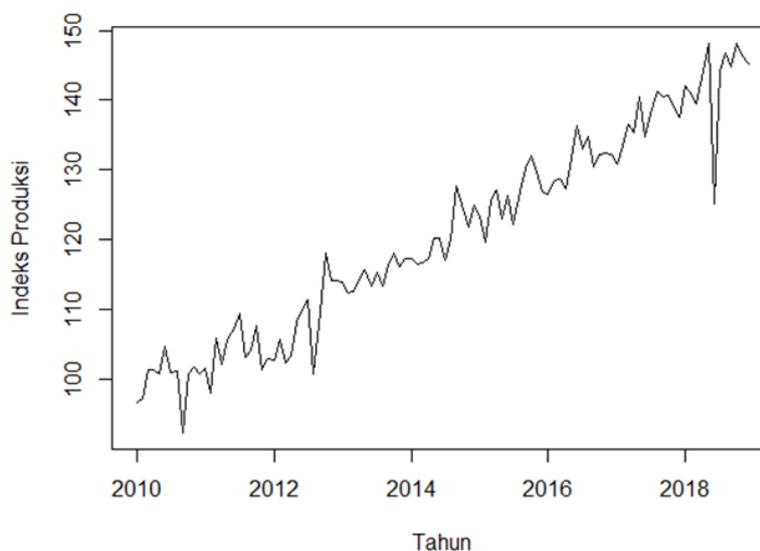
$\hat{f}_{X_{t+b}}$  dan  $\hat{f}_{Y_{t+b}}$  merupakan dugaan kepekatan prediksi pada waktu  $t+b$  untuk *series*  $X_t$  dan  $Y_t$ . Penelitian ini menggunakan  $b=1$  (1 bulan),  $b=3$  (1 triwulan), dan  $b=6$  (1 semester).

6. Melakukan penggerombolan berhirarki dengan metode pautan rataan karena pada berbagai keadaan metode ini lebih stabil dibandingkan dengan pautan tunggal dan pautan lengkap (Mattjik & Sumertajaya, 2011).
7. Menentukan banyaknya gerombol ( $k$ ) optimum dengan melihat nilai rata-rata *silhouette*. Koefisien *silhouette* dihitung untuk 2 sampai 9 gerombol.

### 3. Hasil dan Pembahasan

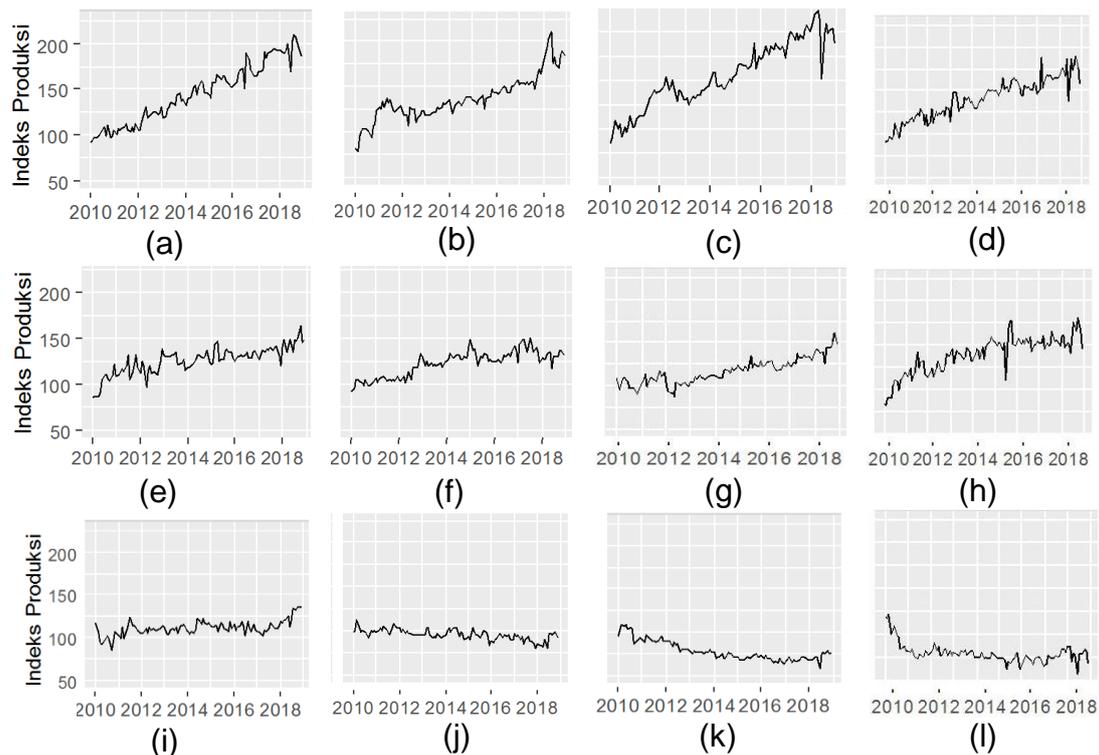
#### 3.1 Deskripsi Data

Dalam kurun waktu tahun 2010 hingga 2018 perkembangan indeks produksi IBS mengalami fluktuasi setiap bulannya, akan tetapi secara rataan tahunan menunjukkan tren meningkat (Gambar 1).



Gambar 1: Perkembangan indeks produksi IBS tahun 2010-2018

Meskipun secara rataan indeks produksi IBS meningkat tetapi masing-masing subsektor memiliki pola perkembangan yang berbeda-beda (Gambar 2). Gambar tersebut memperlihatkan pola perkembangan indeks produksi dari sebagian subsektor. Grafik pola perkembangan indeks produksi menunjukkan 4 subsektor memiliki tren peningkatan yang pesat (KBLI 10, 15, 21, dan 25), 10 subsektor meningkat dengan perlahan (KBLI 12, 14, 18, 20, 23, 24, 27, 28, 29, dan 31), dan 9 subsektor lainnya cenderung stabil bahkan mengalami penurunan.



Gambar 2: Plot perkembangan indeks produksi subsektor IBS KBLI (a) 10, (b) 15, (c) 21, (d) 25, (e) 18, (f) 20, (g) 24, (h) 29, (i) 11, (j) 22, (k) 13, dan (l) 30

### 3.2 Pengujian Kestasioneran Deret Waktu

Pemeriksaan kestasioneran data indeks produksi diperlukan sebelum melakukan analisis gerombol deret waktu. Uji stasioneritas secara formal dapat dilakukan menggunakan uji ADF untuk menguji apakah suatu data deret waktu stasioner dalam rata-rata, serta uji Levene untuk menguji apakah suatu data deret waktu stasioner dalam ragam.

Hasil pengujian menunjukkan hanya ada 7 subsektor yang memenuhi kestasioneran deret waktu baik dalam rata-rata maupun dalam ragam, yaitu industri makanan ( $X_{10}$ ), industri kayu, barang dari kayu dan gabus (tidak termasuk furnitur) dan barang anyaman dari bambu, rotan dan sejenisnya ( $X_{16}$ ), industri kertas dan barang dari kertas ( $X_{17}$ ), industri pencetakan dan reproduksi media rekaman ( $X_{18}$ ), industri karet, barang dari karet dan plastik ( $X_{22}$ ), industri logam dasar ( $X_{24}$ ), serta industri furnitur ( $X_{31}$ ). Sedangkan sisanya tidak stasioner dalam rata-rata dan/atau tidak stasioner dalam ragam. Oleh karena itu, perlu dilakukan transformasi dan pembedaan agar stasioner.

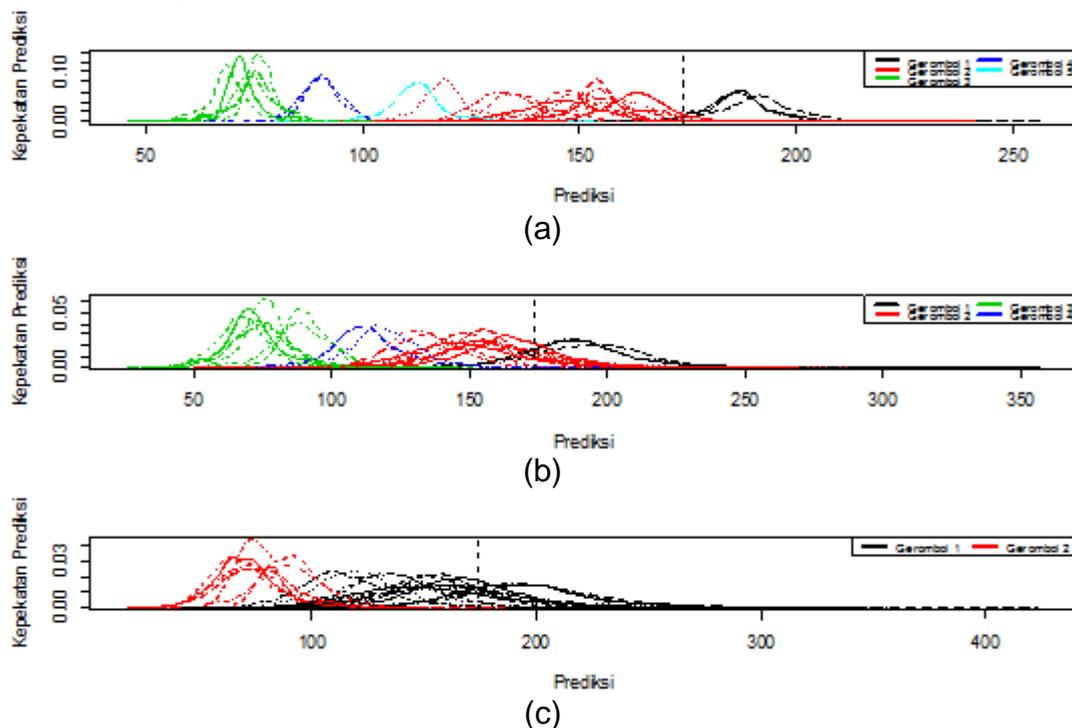
Hasil uji stasioneritas pada data hasil transformasi logaritma dan pembedaan pertama menunjukkan semua subsektor sudah stasioner dalam rata-rata maupun ragam (Tabel 2). Oleh karena itu, proses pembentukan gerombol dapat dilakukan pada data hasil transformasi tersebut. Dengan menggunakan hasil transformasi logaritma natural dan pembedaan 1 tersebut maka penggerombolan didasarkan pada pertumbuhan indeks produksi.

Tabel 2: Hasil uji kestasioneran deret waktu data hasil ransformasi ln dan pembedaan 1 indeks produksi subsektor IBS tahun 2010-2018

Pe- ubah	Nama Peubah	Kestasioneran rataaan		Kestasioneran ragam		Kesimpulan
		Statistik	Nilai	Statistik	Nilai	
		ADF	p	F	p	
X <sub>10</sub>	Industri Makanan	-5.923	0.01	0.040	1.000	stasioner
X <sub>11</sub>	Industri Minuman	-6.414	0.01	0.398	0.997	stasioner
X <sub>12</sub>	Industri Pengolahan Tembakau	-7.011	0.01	0.085	1.000	stasioner
X <sub>13</sub>	Industri Tekstil	-6.154	0.01	0.018	1.000	stasioner
X <sub>14</sub>	Industri Pakaian Jadi	-4.818	0.01	0.066	1.000	stasioner
X <sub>15</sub>	Industri Kulit, Barang dari Kulit dan Alas Kaki	-6.015	0.01	0.111	1.000	stasioner
X <sub>16</sub>	Industri Kayu, Barang dari Kayu dan Gabus (Tidak Termasuk Furnitur) dan Barang Anyaman dari Bambu, Rotan dan Sejenisnya	-5.452	0.01	0.038	1.000	stasioner
X <sub>17</sub>	Industri Kertas dan Barang dari Kertas	-6.713	0.01	0.013	1.000	stasioner
X <sub>18</sub>	Industri Pencetakan dan Reproduksi Media Rekaman	-6.379	0.01	0.095	1.000	stasioner
X <sub>20</sub>	Industri Bahan Kimia dan Barang dari Bahan Kimia	-5.719	0.01	0.076	1.000	stasioner
X <sub>21</sub>	Industri Farmasi, Produk Obat Kimia dan Obat Tradisional	-6.996	0.01	0.076	1.000	stasioner
X <sub>22</sub>	Industri Karet, Barang dari Karet dan Plastik	-5.376	0.01	0.201	1.000	stasioner
X <sub>23</sub>	Industri Barang Galian Bukan Logam	-6.085	0.01	0.047	1.000	stasioner
X <sub>24</sub>	Industri Logam Dasar	-6.71	0.01	0.135	1.000	stasioner
X <sub>25</sub>	Industri Barang Logam, Bukan Mesin dan Peralatannya	-7.839	0.01	0.014	1.000	stasioner
X <sub>26</sub>	Industri Komputer, Barang Elektronik dan Optik	-7.27	0.01	0.165	1.000	stasioner
X <sub>27</sub>	Industri Peralatan Listrik	-6.6	0.01	1.133	0.294	stasioner
X <sub>28</sub>	Industri Mesin dan Perlengkapan ytdl	-5.554	0.01	0.095	1.000	stasioner
X <sub>29</sub>	Industri Kendaraan Bermotor, Trailer dan Semi Trailer	-7.003	0.01	0.028	1.000	stasioner
X <sub>30</sub>	Industri Alat Angkutan Lainnya	-7.418	0.01	0.021	1.000	stasioner
X <sub>31</sub>	Industri Furnitur	-6.689	0.01	0.403	0.996	stasioner
X <sub>32</sub>	Industri Pengolahan Lainnya	-5.51	0.01	0.227	1.000	stasioner
X <sub>33</sub>	Jasa Reparasi dan Pemasangan Mesin dan Peralatan	-6.072	0.01	0.138	1.000	stasioner

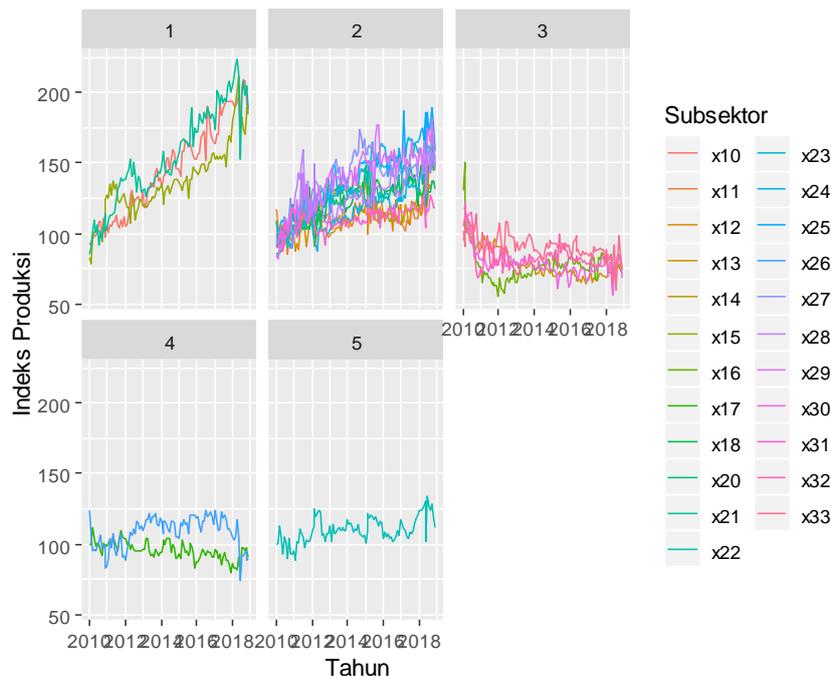
### 3.3 Penggerombolan Untuk Jangka Waktu 1 Bulan

Dengan menggunakan tahun dasar 2010 (2010=100), maka dapat dihitung target indeks IBS pada tahun 2019 sebesar 173.89. Banyaknya gerombol yang terbentuk berdasarkan nilai koefisien *silhouette* yang optimum. Gambar 3 memperlihatkan hasil grafik dugaan kepekatan prediksi menurut gerombol untuk masing-masing jangka waktu prediksi. Sumbu horizontal menunjukkan dugaan nilai prediksi, sumbu vertikal menunjukkan dugaan nilai kepekatan prediksi, sedangkan garis putus-putus hitam menunjukkan target indeks produksi Kemenperin.



Gambar 3: Grafik kepekatan prediksi menurut gerombol untuk jangka waktu prediksi (a) 1 bulan, (b) 1 triwulan, dan (c) 1 semester.

Pada penggerombolan untuk jangka 1 bulan, banyaknya gerombol optimum yang terbentuk adalah 5 gerombol dengan koefisien *silhouette* bernilai 0.4487. Gambar 3(a) menunjukkan gerombol 1 memiliki karakteristik prediksi dengan nilai tengah mendekati 200 dan keragaman yang cukup besar. Gerombol 2 memiliki nilai tengah prediksi sekitar 150 dengan keragaman yang juga cukup besar, sedangkan subsektor pada gerombol 3 diprediksikan dibawah 100 dengan keragaman yang kecil. Gambar 3(a) juga memperlihatkan gerombol 1 memiliki prediksi yang melebihi target Kemenperin, sedangkan gerombol lainnya tidak memenuhi target, sehingga untuk jangka waktu 1 bulan ke depan subsektor pada gerombol 1 layak dipilih menjadi industri andalan di masa mendatang. Subsektor pada gerombol 1 yaitu industri makanan ( $X_{10}$ ); industri kulit, barang dari kulit dan alas kaki ( $X_{15}$ ) dan industri farmasi, produk obat kimia dan obat tradisional ( $X_{21}$ ).



Gambar 4: Pola perkembangan indeks produksi menurut hasil gerombol *prediction based* untuk jangka waktu 1 bulan.

Gambar 4 menunjukkan pola perkembangan indeks produksi pada masing-masing gerombol. Gambar tersebut memperlihatkan gerombol 1 memiliki tren pola perkembangan indeks produksi yang sangat pesat, sehingga diprediksi memiliki nilai tengah yang paling tinggi dibandingkan gerombol yang lain. Gerombol 2 memiliki pola perkembangan yang cukup pesat namun tidak lebih tinggi dari gerombol 1. Sedangkan pada gerombol 3, pola perkembangan indeks produksi mengalami penurunan dan cenderung menuju stabil pada nilai di bawah 100. Gerombol 4 terlihat tidak memiliki tren dan cenderung berpola tidak linier. Sedangkan gerombol 5 berfluktuasi dan memiliki tren meningkat tetapi tidak sebesar gerombol 2.

### 3.4 Penggerombolan Untuk Jangka Waktu 1 Triwulan

Pada penggerombolan untuk jangka 1 triwulan, banyaknya gerombol optimum yang terbentuk adalah 4 gerombol dengan koefisien silhouette bernilai 0.5477. Gambar 3(b) menunjukkan karakteristik hasil penggerombolan berdasarkan fungsi kepekatan prediksi untuk jangka waktu 1 triwulan. Gambar tersebut memperlihatkan gerombol 1 secara konsisten memiliki nilai tengah prediksi yang paling tinggi dan juga ragam yang besar, sedangkan gerombol 3 juga konsisten diprediksi memiliki nilai tengah di bawah 100 dan keragaman yang kecil. Gambar tersebut juga memperlihatkan gerombol 1 memiliki prediksi yang melebihi target kemenperin, sedangkan gerombol lainnya tidak memenuhi target, sehingga untuk jangka 1 triwulan subsektor pada gerombol 1 juga layak dipilih menjadi industri andalan di masa mendatang. Subsektor pada gerombol 1 untuk jangka 1 triwulan ke depan sama dengan subsektor pada gerombol 1 untuk jangka 1 bulan ke depan.

Pola perkembangan indeks produksi masing-masing gerombol berdasarkan penggerombolan *prediction based* untuk jangka 1 triwulan memiliki karakteristik yang

lebih jelas (Gambar 5). Gambar tersebut memperlihatkan ada 4 pola perkembangan yang berbeda. Gerombol 1 memiliki pola perkembangan indeks produksi yang sangat pesat, diikuti oleh gerombol 2 yang memiliki pola perkembangan yang cukup pesat, lalu disusul gerombol 4 yang memiliki pola perkembangan positif tetapi tidak pesat. Sedangkan gerombol 3 memiliki pola perkembangan yang menurun atau cenderung menuju stabil pada nilai rata-rata di bawah 100.

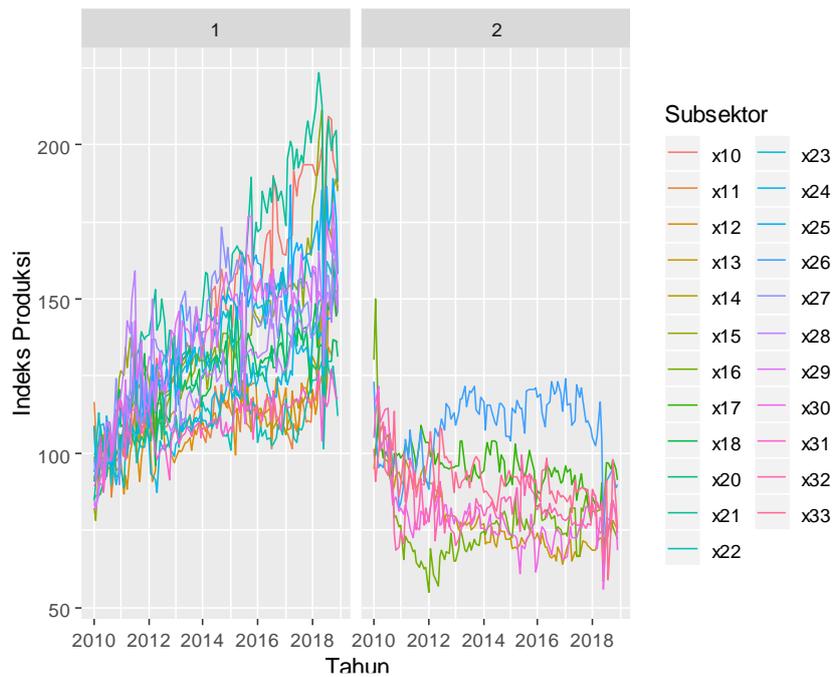


Gambar 5: Pola perkembangan indeks produksi menurut hasil gerombol *prediction based* untuk jangka waktu 1 triwulan.

### 3.5 Penggerombolan Untuk Jangka Waktu 1 Semester

Pada penggerombolan untuk jangka 1 semester, banyaknya gerombol optimum yang terbentuk adalah 2 gerombol dengan koefisien silhouette bernilai 0.5645. Terdapat 2 karakteristik prediksi yang berbeda, yaitu gerombol 1 memiliki nilai tengah prediksi di atas 100 dengan keragaman yang besar, sedangkan gerombol 2 memiliki nilai tengah prediksi di bawah 100 dengan keragaman yang kecil (Gambar 3c). Gambar tersebut juga memperlihatkan gerombol 1 terdiri atas campuran subsektor yang mampu memenuhi target dan yang tidak dapat memenuhi target Kemenperin.

Karena hanya 2 gerombol yang terbentuk maka karakteristik pola perkembangan indeks produksi pada Januari 2010 - Desember 2018 terlihat sangat berbeda. Subsektor pada gerombol 1 memiliki pola perkembangan positif, sedangkan subsektor pada gerombol 2 memiliki pola perkembangan tidak positif atau menurun dan cenderung menuju stabil (Gambar 6).



Gambar 6: Pola perkembangan indeks produksi menurut hasil gerombol *prediction based* untuk jangka waktu 1 semester.

### 3.6 Perbandingan Hasil Penggerombolan

Terdapat perbedaan karakteristik dari hasil penggerombolan untuk jangka waktu prediksi 1 bulan, 1 triwulan, dan 1 semester. Dilihat dari grafik fungsi kepekatan prediksi pada Gambar 3, grafik fungsi kepekatan prediksi untuk jangka 1 bulan memiliki puncak yang lebih curam dibandingkan dengan jangka 1 triwulan dan 1 semester. Dengan demikian semakin panjang jangka waktu prediksi semakin besar keragaman atau semakin tidak pasti nilai prediksinya.

Tabel 3: Perbandingan rata-rata koefisien *silhouette*

Banyaknya Gerombol	Jangka Waktu Prediksi		
	$b=1$	$b=3$	$b=6$
2	0.2725	0.4603	0.5645
3	0.2567	0.4465	0.4905
4	0.3621	0.5477	0.5662
5	0.4487	0.5727	0.6063
6	0.4478	0.5870	0.6457
7	0.4434	0.5932	0.5916
8	0.4766	0.5957	0.4835
9	0.5409	0.4962	0.4952
Rata-rata	0.4061	0.5374	0.5554

Tabel 3 menyajikan perbandingan rata-rata koefisien *silhouette* untuk berbagai jangka waktu prediksi (1 bulan, 1 triwulan, dan 1 semester). Secara rata-rata semakin panjang jangka waktu prediksi rata-rata koefisien *silhouette* semakin tinggi. Hal ini menunjukkan semakin panjang jangka waktu prediksi hasil penggerombolan semakin

baik, sehingga penggerombolan dengan *prediction-based* yang terbaik adalah untuk jangka prediksi 1 semester. Akan tetapi, hasil penggerombolan optimum untuk jangka 1 semester pada gerombol 1 terdiri atas campuran subsektor yang mampu memenuhi target dan yang tidak dapat memenuhi target Kemenperin. Dengan hasil tersebut, tidak dapat secara eksplisit menentukan industri mana yang akan menjadi industri andalan masa depan. Oleh karena itu, dipilih penggerombolan terbaik adalah untuk jangka waktu prediksi 1 triwulan.

Dilihat dari karakteristik masing-masing gerombol pada periode waktu penelitian yaitu Januari 2010- Desember 2018, grafik perkembangan indeks produksi menurut gerombol untuk jangka waktu prediksi 1 triwulan menghasilkan pola yang cukup mewakili masing-masing karakteristik gerombol, yaitu pola perkembangan indeks produksi yang sangat pesat, pola perkembangan yang cukup pesat, pola perkembangan positif tetapi tidak pesat serta pola perkembangan yang menurun atau cenderung menuju stabil pada nilai rata-rata di bawah 100 (Gambar 5). Subsektor hasil penggerombolan untuk jangka waktu 1 triwulan disajikan pada Tabel 4.

Tabel 4: Hasil penggerombolan *prediction based* untuk 1 triwulan ke depan

Gerombol	Subsektor
1	Industri Makanan; Industri Kulit, Barang dari Kulit dan Alas Kaki; Industri Farmasi, Produk Obat Kimia dan Obat Tradisional
2	Industri Minuman; Industri Pengolahan Tembakau; Industri Pakaian Jadi; Industri Pencetakan dan Reproduksi Media Rekaman; Industri Bahan Kimia dan Barang dari Bahan Kimia; Industri Barang Galian Bukan Logam; Industri Logam Dasar; Industri Barang Logam, Bukan Mesin dan Peralatannya; Industri Peralatan Listrik; Industri Mesin dan Perlengkapan ytdl; Industri Kendaraan Bermotor, Trailer dan Semi Trailer
3	Industri Tekstil; Industri Kayu, Barang dari Kayu dan Gabus (Tidak Termasuk Furnitur) dan Barang Anyaman dari Bambu, Rotan dan Sejenisnya; Industri Alat Angkutan Lainnya; Industri Pengolahan Lainnya; Jasa Reparasi dan Pemasangan Mesin dan Peralatan; Industri Kertas dan Barang dari Kertas; Industri Komputer, Barang Elektronik dan Optik
4	Industri Furnitur; Industri Karet, Barang dari Karet dan Plastik

#### 4. Simpulan dan Saran

Perkembangan indeks produksi IBS masing-masing subsektor memiliki pola yang berbeda-beda. Dari hasil penggerombolan *prediction-based* dapat disimpulkan bahwa industri makanan ( $X_{10}$ ); industri kulit, barang dari kulit dan alas kaki ( $X_{15}$ ) dan industri farmasi, produk obat kimia dan obat tradisional ( $X_{21}$ ) dapat dipilih untuk dijadikan sebagai sektor industri andalan masa depan karena ketiganya mampu memenuhi target pertumbuhan indeks produksi yang diberikan oleh Kemenperin di masa mendatang.

**Daftar Pustaka**

- Alonso, A. M., Berrendero, J. R., Hernández, A., & Justel, A. (2006). Time series clustering based on forecast densities. *Computational Statistics & Data Analysis*, 51(2): 762–776.
- Hardle, W. (1994). *Applied Nonparametric Regression*. New York (US): Cambridge University Press.
- Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern Recognition*, 38(11): 1857–1874.
- Mattjik, A. A., & Sumertajaya, I. M. (2011). *Sidik Peubah Ganda Dengan Menggunakan SAS*. Bogor (ID): IPB Press.
- Montero, P., & Vilar, J. A. (2014). TSclust: An R Package for Time Series Clustering. *Journal of Statistical Software*, 62(1): 1–43.
- Sheather, S. J., & Jones, M. C. (1991). A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. *Journal of the Royal Statistical Society*, 53(3): 683–690.
- Sukarsa, I. K. G., & Srinadi, I. G. A. M. (2012). ESTIMATOR KERNEL DALAM MODEL REGRESI NONPARAMETRIK. *Jurnal Matematika*, 2(1): 19–30.
- Vilar, J. A., Alonso, A. M., & Vilar, J. M. (2010). Non-linear time series clustering based on non-parametric forecast densities. *Computational Statistics & Data Analysis*, 54(11): 2850–2865.