

PROFIL PROVINSI JAWA BARAT BERDASARKAN ANGKA PARTISIPASI SEKOLAH MENGGUNAKAN *PRINCIPAL COMPONENT ANALYSIS BIPLLOT*

TITI PURWANDARI¹, SOEMARTINI², NENENG SUNENSIH³

¹Departemen Statistika FMIPA Universitas Padjadjaran, titipurwandari@yahoo.com

²Departemen Statistika FMIPA Universitas Padjadjaran, tine_soemartini@yahoo.com

³Departemen Statistika FMIPA Universitas Padjadjaran, nenks.stat@gmail.com

Abstrak. Bersekolah adalah mereka yang terdaftar dan aktif mengikuti pendidikan baik jenjang pendidikan formal berupa pendidikan dasar yaitu Sekolah Dasar/ sederajat dan Sekolah Menengah Pertama/ sederajat, pendidikan menengah yaitu Sekolah Menengah Atas/ sederajat dan pendidikan tinggi yaitu Perguruan Tinggi/ sederajat maupun non formal Paket A setara SD, paket B setara SMP dan paket C setara SMA yang berada di bawah pengawasan Kementerian Pendidikan Nasional (Kemdiknas), Sekolah (APS) merupakan proporsi anak sekolah pada usia jenjang pendidikan tertentu dalam kelompok usia yang sesuai dengan jejang pendidikan tersebut. Gejala tingkat pendidikan yang rendah mencirikan bahwa investasi sumber daya manusia belum dilakukan secara memadai. Tujuan penelitian ini adalah mengetahui profil provinsi Jawa Barat dengan cara mengelompokkan kabupaten dan kota di Jawa Barat berdasarkan Angka Partisipasi Sekolah dalam rangka memberi rekomendasi kepada instansi terkait. Kegunaan dari penelitian ini adalah memberi referensi ilmiah bagi pemerintah Jawa Barat dalam membuat kebijakan. Data yang digunakan merupakan data sekunder yang dikumpulkan Badan Pusat Statistik. Hasil analisis menggunakan *Principal Component Analysis Biplot*, diperoleh peta pengelompokan kemiripan kabupaten dan kota berdasarkan angka partisipasi sekolah, keragaman angka partisipasi sekolah, korelasi angka partisipasi sekolah.

Kata kunci: Angka Partisipasi Sekolah, Peta Pengelompokan, *Principal Component Analysis Biplot*.

1. Pendahuluan

Bersekolah adalah mereka yang terdaftar dan aktif mengikuti pendidikan baik jenjang pendidikan formal berupa pendidikan dasar yaitu Sekolah Dasar/ sederajat dan Sekolah Menengah Pertama/ sederajat, pendidikan menengah yaitu Sekolah Menengah Atas/ sederajat dan pendidikan tinggi yaitu Perguruan Tinggi/ sederajat) maupun non formal Paket A setara SD, paket B setara SMP dan paket C setara SMA) yang berada di bawah pengawasan Kementerian Pendidikan Nasional (Kemdiknas), Sekolah (APS) merupakan proporsi anak sekolah pada usia jenjang pendidikan tertentu dalam kelompok usia yang sesuai dengan jejang pendidikan tersebut (<https://www.bps.go.id>).

Pendidikan formal adalah jalur pendidikan yang terstruktur dan berjenjang yang terdiri atas pendidikan dasar, pendidikan menengah, dan pendidikan tinggi, meliputi SD/MI/ sederajat, SMP/MTs/ sederajat, SM/MA/ sederajat dan PT. Pendidikan nonformal adalah jalur pendidikan di luar pendidikan formal yang dapat dilaksanakan secara terstruktur dan berjenjang. Meliputi pendidikan kecakapan hidup (kursus), pendidikan anak usia dini (PAUD) atau pra-sekolah, pendidikan kepemudaan, pendidikan pemberdayaan perempuan, pendidikan keaksaraan, pendidikan keterampilan dan pelatihan kerja, pendidikan kesetaraan (paket A, paket B, dan paket C) serta pendidikan lainnya yang ditujukan untuk mengembangkan kemampuan peserta

didik. Tidak/belum pernah sekolah adalah tidak/belum pernah terdaftar dan aktif mengikuti pendidikan di suatu jenjang pendidikan, termasuk mereka yang tamat/belum tamat Taman Kanak-kanak yang tidak melanjutkan ke Sekolah Dasar. Tamat sekolah adalah telah menyelesaikan pelajaran pada kelas/tingkat terakhir suatu jenjang pendidikan di sekolah negeri maupun swasta dengan mendapatkan tanda tamat/ijazah. Seorang yang belum mengikuti pelajaran pada kelas tertinggi tetapi jika ia mengikuti ujian dan lulus maka dianggap tamat. Angka Buta Huruf adalah proporsi penduduk usia tertentu yang tidak dapat membaca dan atau menulis huruf Latin atau huruf lainnya terhadap penduduk usia tertentu. Gejala tingkat pendidikan yang rendah mencirikan bahwa investasi sumber daya manusia belum dilakukan secara memadai [4].

Tujuan penelitian ini adalah mengetahui profil provinsi Jawa Barat dengan cara mengelompokkan kabupaten dan kota di Jawa Barat berdasarkan Angka Partisipasi Sekolah dalam rangka memberi rekomendasi kepada instansi terkait. Kegunaan dari penelitian ini adalah memberi referensi ilmiah bagi pemerintah Jawa Barat dalam membuat kebijakan.

Sistematika pengelompokan kabupaten dan kota di Jawa Barat berdasarkan Angka Partisipasi Sekolah dianalisis menggunakan metoda *Principal Component Analysis Biplot (PCA) Biplot* yang merupakan bagian dari analisis multivariat. *PCA* merupakan metode pemetaan dalam analisis multivariat yang memuat informasi dalam sebuah tabel data, yang menunjukkan struktur utama data. Analisis ini bertujuan untuk menyajikan data dalam peta dua dimensi sehingga perilaku data mudah dilihat dan diinterpretasikan.

2. Metode

2.1 Obyek dan Variabel Penelitian

Obyek pengamatan pada penelitian ini adalah kabupaten dan kota di Provinsi Jawa Barat yaitu sebanyak 26 kabupaten dan kota yang terdiri dari tujuh belas kabupaten dan Sembilan kota. Variabel dalam penelitian ini adalah variabel angka partisipasi sekolah umur 7-12 tahun, 13-15 tahun, 16-18 tahun, dan 19-24 tahun. Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik tahun 2015 [4].

2.2 *Principal Component Analysis Biplot (PCA) Biplot*

Principal Component Analysis Biplot merupakan metode pemetaan dalam analisis multivariat yang memuat informasi dalam sebuah tabel data, yang menunjukkan struktur utama data[2]. Analisis ini bertujuan untuk menyajikan data dalam peta dua dimensi sehingga perilaku data mudah dilihat dan diinterpretasikan. Analisis biplot memerlukan data dari sejumlah objek dengan variabel berskala interval atau rasio. Metode ini didasarkan pada *Singular Value Decomposition (SVD)* dari suatu matriks data yang telah terkoreksi oleh rata rata. Langkah langkah analisis ini adalah sebagai berikut :

2.2.1 Matriks Data

Data berupa objek sebanyak 27 kabupaten-kota dengan variabel sebanyak 4 disajikan ke dalam matriks awal \mathbf{Y} yang berukuran $n \times p$ (27×4).

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{14} \\ y_{21} & y_{22} & \dots & y_{24} \\ \vdots & \vdots & \ddots & \vdots \\ y_{271} & y_{272} & \dots & y_{274} \end{bmatrix}$$

Matriks \mathbf{Y} ditransformasi terhadap rata-rata.

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} y_{11} - \bar{y}_1 & y_{12} - \bar{y}_2 & \dots & y_{14} - \bar{y}_4 \\ y_{21} - \bar{y}_1 & y_{22} - \bar{y}_2 & \dots & y_{24} - \bar{y}_4 \\ \vdots & \vdots & \ddots & \vdots \\ y_{27.1} - \bar{y}_1 & y_{27.2} - \bar{y}_2 & \dots & y_{274} - \bar{y}_4 \end{bmatrix} \\ &= \begin{bmatrix} x_{11} & x_{12} & \dots & x_{14} \\ x_{21} & x_{22} & \dots & x_{24} \\ \vdots & \vdots & \ddots & \vdots \\ x_{271} & x_{272} & \dots & x_{274} \end{bmatrix} \end{aligned}$$

2.2.2 Eigenvalue dan Eigenvector

Sebelum mencari nilai dekomposisi singularnya (SVD) perlu dihitung *eigenvalue* dan *eigenvector* dari matriks data $\mathbf{X}^T \mathbf{X}$ [5]. *Eigenvalue* dinotasikan dengan λ dan *eigenvector* yang dinotasikan dengan \mathbf{a} dapat dihitung sebagai berikut :

$$\begin{aligned} |\mathbf{X}^T \mathbf{X} - \lambda_i \mathbf{I}| &= 0 \\ (\mathbf{X}^T \mathbf{X} - \lambda_i) \mathbf{a} &= 0 \end{aligned}$$

2.2.3 Singular Value Decomposition (SVD)

Pendekatan langsung untuk memperoleh nilai dekomposisi singularnya (SVD) adalah sebagai berikut [5]:

$$\mathbf{X}_{(n \times p)} = \mathbf{U}_{(n \times r)} \mathbf{L}_{(r \times r)} \mathbf{A}^T_{(r \times p)}$$

dengan :

$$r \leq \{27, 4\}$$

- \mathbf{U} dan \mathbf{A}^T adalah matriks dengan kolom ortonormal sehingga $\mathbf{U}^T \mathbf{U} = \mathbf{A}^T \mathbf{A} = \mathbf{I}_r$ (\mathbf{I}_r adalah matriks identitas berdimensi r)
- \mathbf{L} adalah matriks berukuran ($r \times r$) dengan unsur-unsur diagonalnya adalah akar kuadrat dari *eigenvalue* $\mathbf{X}^T \mathbf{X}$, dengan $\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_r}$ yang membentuk matriks sebagai berikut :

$$\mathbf{L} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{\lambda_r} \end{bmatrix}$$

- Unsur-unsur diagonal matriks disebut dengan nilai singular matriks \mathbf{X} .
- Kolom-kolom matriks \mathbf{A} adalah *eigenvector* dari matriks $\mathbf{X}^T \mathbf{X}$ yang berpadanan dengan *eigenvalue* λ_i yaitu :

$$\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r]$$

- Kolom-kolom matriks \mathbf{U} didapatkan dari rumus :

$$\mathbf{u}_i = \frac{1}{\sqrt{\lambda_i}} \mathbf{X}\mathbf{a}_i, \quad i = 1, 2, \dots, r$$

dengan:

- \mathbf{u}_i : unsur-unsur matriks \mathbf{U}
 - \mathbf{a}_i : unsur-unsur matriks \mathbf{A}
 - λ_i : *eigenvalue* ke- i dari matriks $\mathbf{X}^T\mathbf{X}$
 - \mathbf{X} : matriks asal yang dikoreksi terhadap rata-rata
- Setelah hasil *SVD* didapatkan, maka menjadi persamaan :

$$\mathbf{X} = \mathbf{U}\mathbf{L}^\alpha\mathbf{L}^{1-\alpha}\mathbf{A}^T$$

Dalam menentukan \mathbf{L}^α , untuk $0 \leq \alpha \leq 1$, maka matriks diagonalnya mempunyai unsur-unsur diagonalnya $\sqrt{\lambda_1^\alpha} \geq \sqrt{\lambda_2^\alpha} \geq \dots \geq \sqrt{\lambda_r^\alpha}$. Penentuan $\mathbf{L}^{1-\alpha}$ berlaku sama dengan unsur-unsur diagonalnya $\sqrt{\lambda_1^{1-\alpha}} \geq \sqrt{\lambda_2^{1-\alpha}} \geq \dots \geq \sqrt{\lambda_r^{1-\alpha}}$.

Misalkan $\mathbf{G} = \mathbf{U}\mathbf{L}^\alpha$ dan $\mathbf{H}^T = \mathbf{L}^{1-\alpha}\mathbf{A}^T$

$$\mathbf{G}\mathbf{H}^T = \mathbf{U}\mathbf{L}^\alpha\mathbf{L}^{1-\alpha}\mathbf{A}^T = \mathbf{U}\mathbf{L}\mathbf{A}^T = \mathbf{X}$$

Unsur ke- (i,j) dalam matriks \mathbf{X} dapat dituliskan sebagai berikut :

$$\mathbf{x}_{ij} = \mathbf{g}_i\mathbf{h}_j$$

Dengan \mathbf{g}_i , $i = 1, 2, \dots, 27$ dan \mathbf{h}_j , $j = 1, 2, 3, 4$ masing-masing merupakan baris matriks \mathbf{G} dan kolom matriks \mathbf{H} . Pada \mathbf{g}_i dan \mathbf{h}_j mempunyai r dimensi. Dua kolom pertama dari matriks \mathbf{G} dapat digunakan untuk pemetaan objek, sedangkan dua kolom pertama dari matriks \mathbf{H}^T digunakan untuk pemetaan variabel.

2.2.4 Identifikasi Persentase Keragaman Data

Jika matriks \mathbf{X} mempunyai rank lebih dari dua maka *eigenvalue* yang diambil adalah λ_1 dan λ_2 sehingga besarnya keragaman yang diterangkan adalah sebagai berikut :

$$\tau = \frac{(\lambda_1 + \lambda_2)}{\sum_{i=1}^p \lambda_i}$$

- λ_1 : *Eigenvalue* terbesar pertama
- λ_2 : *Eigenvalue* terbesar kedua
- λ_i : *Eigenvalue* ke- i dari $\mathbf{X}^T\mathbf{X}$; $i = 1, 2, \dots, 27$

Jika nilai τ semakin mendekati nilai 1 berarti biplot yang diperoleh dari matriks pendekatan dengan rank = 2 akan memberikan penyajian yang semakin baik mengenai informasi-informasi yang terdapat pada data yang sebenarnya. Maka berdasarkan nilai τ tersebut peta pengelompokan yang dihasilkan dapat digunakan dalam pengambilan keputusan. Peta pengelompokan yang dihasilkan dapat memberikan gambaran posisi kedekatan satu objek dengan objek lainnya serta kedekatan variabel terhadap objek.

2.2.5 Identifikasi Hasil Pemetaan *Principal Component Analysis Biplot* (Biplot PCA)

Hasil pemetaan dari Biplot *PCA* adalah sebagai berikut [2] :

1. Kedekatan (similaritas) antar objek penelitian.
Semakin dekat posisi dua buah titik objek maka semakin mirip ,semakin jauh posisi dua buah titik objek maka semakin berbeda .
2. Keragaman variabel.
Variabel digambarkan sebagai garis berarah (vektor). Variabel dengan keragaman kecil digambarkan sebagai vektor berukuran pendek sedangkan variabel dengan keragaman besar digambarkan sebagai vektor berukuran panjang.
3. Hubungan atau korelasi antar variabel.
Hubungan antar variabel dapat diidentifikasi berdasarkan sudut yang terbentuk dari dua buah vektor variabel terhadap sumbu peta. Jika dua buah vektor variabel berhimpit terhadap sumbu peta dengan arah yang sama (mendekati 0° atau 360°), maka memiliki korelasi positif sangat erat. Jika dua buah vektor variabel berhimpit dengan sumbu peta dengan arah berlawanan (mendekati 180°), maka memiliki korelasi negatif yang sangat erat, jika dua buah vektor variabel tegak lurus terhadap sumbu peta (mendekati 90° atau 270°) maka dua buah variabel tersebut tidak berkorelasi.
4. Nilai variabel pada suatu objek.
Objek yang terletak searah dengan arah dari suatu variabel, dikatakan bahwa pada objek tersebut nilainya di atas rata-rata. Sebaliknya, jika objek lain terletak berlawanan dengan arah dari variabel tersebut, maka objek tersebut memiliki nilai di bawah rata-rata. Sedangkan objek yang hampir di tengah-tengah, memiliki nilai dekat dengan rata-rata.

3. Hasil dan Pembahasan

3.1 Identifikasi Persentase Keragaman Data

Tabel 1 menunjukkan jumlah komponen yang terbentuk yaitu 2 komponen dan Tabel 2 menunjukkan besar keragaman data yang dapat dijelaskan oleh peta PCA Biplot menggunakan 2 komponen dihitung dari besar kumulatif eigenvalue adalah 0,83020 , maka dapat disimpulkan bahwa persentase keragaman data yang dapat dijelaskan oleh PCA Biplot adalah sebesar 83,020 %. Hal ini dapat dilihat pada tabel sebagai berikut :

Tabel 1 Jumlah Komponen Yang Terbentuk

Component Matrix ^a		
	Component	
	1	2
umur 7-12	.006	.987
umur 13-15	.827	.221
umur 16-18	.861	.007
umur 19-24	.909	-.214

Extraction Method: Principal Component Analysis.

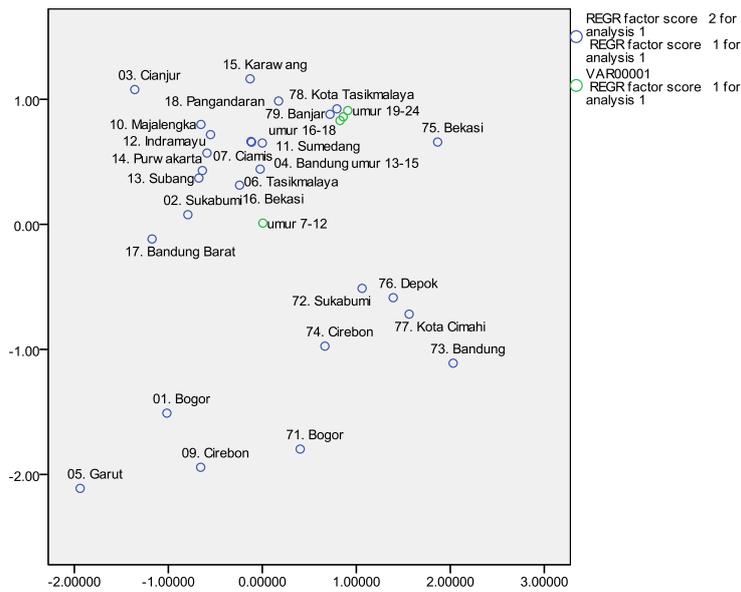
a. 2 components extracted.

Tabel 2 Keragaman Data Yang Dijelaskan oleh PCA Biplot

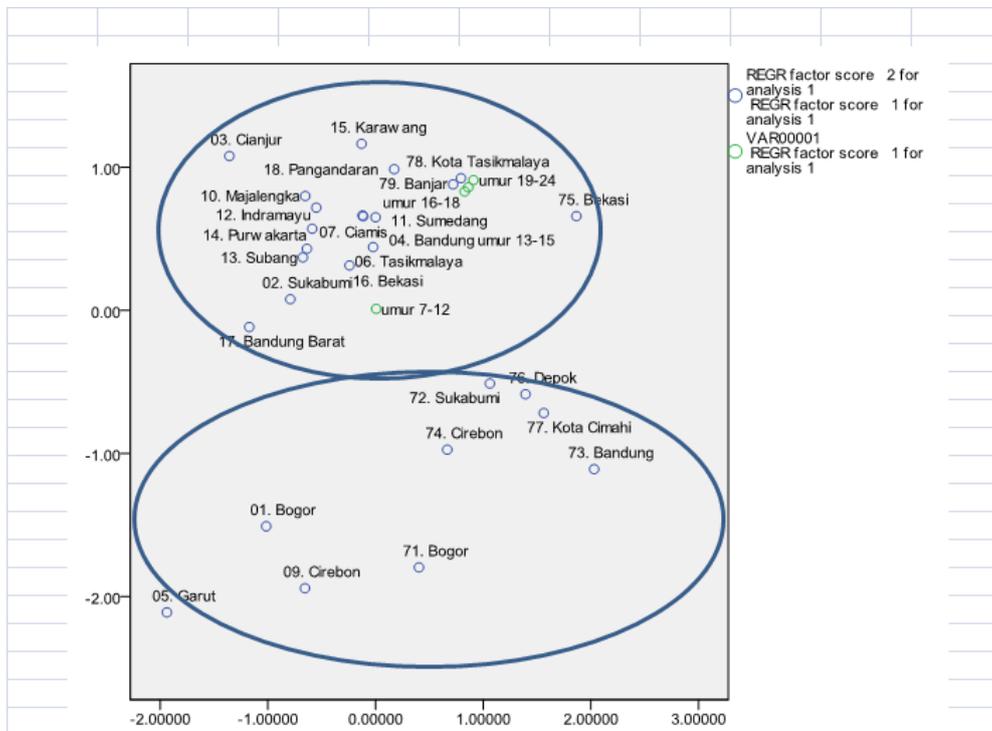
Compo nent	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.252	56.288	56.288
2	1.069	26.732	83.020
3	.461	11.531	94.552
4	.218	5.448	100.000

3.2 Identifikasi informasi Hasil Analisis Pemetaan

Grafik 1 menunjukkan hasil analisis melalui peta , terbentuk 2 buah kelompok kabupaten/kota dengan karakteristik variabel memiliki kemiripan. Kelompok 1 terdiri dari 18 kabupaten dan kota yaitu kabupaten kabupaten Cianjur, kab Karawang, Pangandaran, Majalengka, Indramayu, Purwakarta, Ciamis, Subang, Sukabumi, Sumedang, Bandung, Tasikmalaya, Bandung Barat, Kuningan, Bekasi, kota Tasikmalaya, kota Banjar, kota Bekasi. Kelompok 2 terdiri dari 9 kabupaten dan kota yaitu kota kota Depok, Sukabumi, Cirebon, Cimahi, Bandung, Bogor, kab Bogor, kab Cirebon, kab Garut.



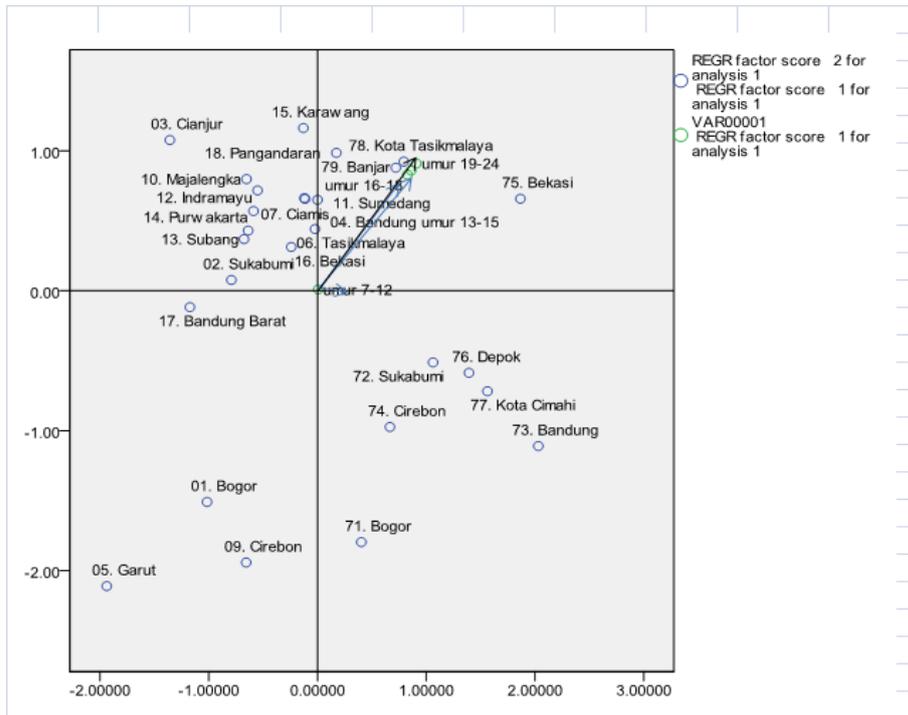
Grafik 1 Pemetaan Kabupaten dan Kota di Jawa Barat



Grafik 2 Peta Pengelompokan Kabupaten Kota

3.3 Keeratan Hubungan Antar Variabel Penelitian

Grafik 3 menunjukkan Keeratan hubungan antar variabel yang diamati berdasarkan pengelompokan kabupaten dan kota di Jawa Barat. Dari Grafik tersebut dapat dilihat bahwa angka partisipasi sekolah umur 13-15 tahun, umur 16-18 tahun, dan umur 19-24 tahun memiliki hubungan yang sangat erat, hal ini ditunjukkan melalui sudut yang dibentuk oleh kedua garis yang menunjukkan angka partisipasi sekolah.



Grafik 3 Peta Hubungan Antar Variabel

3. Kesimpulan

Berdasarkan hasil analisis menggunakan Principal Component Analysis Biplot diperoleh kesimpulan bahwa terbentuk 2 buah kelompok kabupaten/kota dengan karakteristik variabel memiliki kemiripan, persentase keragaman data yang dapat dijelaskan oleh PCA Biplot adalah sebesar 83,020 %, angka partisipasi sekolah umur 13-15 tahun, umur 16-18 tahun, dan umur 19-24 tahun memiliki hubungan yang sangat erat, hal ini ditunjukkan melalui sudut yang dibentuk oleh kedua garis yang menunjukkan angka partisipasi sekolah, Variabel yang memiliki keragaman terbesar adalah angka partisipasi sekolah umur 19-24 tahun, sedangkan Variabel yang memiliki keragaman terkecil adalah angka partisipasi sekolah umur 7-12 tahun.

Referensi

- [1] Gabriel, K. R. (1971). *The Biplot Graphic Display of Matrices with Application to Principal Component Analysis*. Biometrika 58-45.
- [2] Grenaacre, M. J. (2010). *Biplots in Practice*. Barcelona : BBVA Foundation, The Pompeu Fabra University Barcelona.
- [3] Hair ,J.F., Anderson , R.E.,Tatham, R.L.,and Black,Inc., W.C. (2010). *Multivariate Data Analysis*, Seven Edition, Prentice Hall International,Upper Saddle River, New Jersey.
- [4] <https://www.bps.go.id>, Badan Pusat Statistik.
- [5] Johnson, Wichern. (2007). *Applied Multivariate Statistical Analysis*, Sixth Edition. Prentice Hall International, Inc., Upper Saddle River, New Jersey.