

MENGGUNAKAN DAN MENGINTERPRETASIKAN EXPLORATION FACTOR ANALYSIS: PANDUAN BAGI PENELITI DAN MAHASISWA

Adnan Sauddin

Dosen Jurusan Matematika, Fak. Sains dan Teknologi.

UIN Alauddin Makassar

Adnan.sauddin@uin-alauddin.ac.id

Abstract: *How should we draw the relationship between each variable in a study included a number of variables, especially research that is very much. A simple question, but it will have a profound effect on the validity of the research is not done. Factor analysis is a statistical analysis technique that can help conduct research on a set of variables in the grouping factors, where each variable in the variable has a large variance. The use of statistical science as a tool to analyze, especially if the use of statistical science aided by using software such as Minitab, SPSS, SAS. That uses software to calculate statistics, an expression of "whatever we put in a certain statistical software will still produce something in the form of numbers", the question is whether the resulting output is valid or not? These are important questions. Therefore, in this article provided a good explanation of the theory of mathematical statistics, along with the use of software (Minitab) to facilitate the calculation process, also given the interpretation of the results obtained.*

Key words: *Analisis Faktor, minitab.*

I. PENGANTAR

Analisis factor adalah suatu teknik inter-dependensi (kebergantungan antara) yang tujuan utamanya adalah untuk mendefinisikan struktur yang diperhatikan antara variable dalam analisis. Variable merupakan kunci utama dalam analisis multivariate. Apakah kita membuat suatu peramalan penjualan dengan analisis regresi, melakukan prediksi terhadap gagal atau suksesnya suatu perusahaan baru dengan analisis diskriminan, atau menggunakan teknik multivariate yang lain, kita harus mempunyai sekumpulan variable yang membentuk hubungan.

Ketika kita melakukan penelitian terhadap sekumpulan variable, analisis univariat tidak relevan untuk kita gunakan. Untuk kondisi seperti itu, yaitu jumlah variable yang diikuti dalam suatu penelitian banyak, maka analisis multivariat sangat pantas untuk digunakan untuk menjawab berbagai pertanyaan yang berkaitan dengan keadaan dari sekumpulan variable tersebut. Pertanyaannya adalah bagaimana kita menjelaskan dan menggambarkan seluruh variable tersebut? Tentunya, jika kita hanya sedikit variable yang kita ikuti dalam suatu penelitian, hal tersebut mungkin berbeda! Akan tetapi, ketika kita tambahkan dan menambahkan lagi variable yang lebih banyak, maka kemungkinan saling overlap atau tumpang tindih atau saling berkorelasi antara satu variable dengan variable yang lain sangatlah tinggi. Sebagai contoh, kita menggunakan berbagai ukuran terhadap variable yang banyak maka kemungkinan kesalahan pengukuran akan sangat banyak terjadi, pada saat yang sama peneliti ingin mengetahui korelasi antara setiap variable. Ketika variable-variable tersebut saling berkorelasi, maka peneliti memerlukan suatu cara mengatasi keadaan variable-variable tersebut dengan cara-mengumpulkan variable-variable dengan korelasi yang tinggi dalam satu kelompok, memberikan nama terhadap kelompok-kelompok tersebut, dan mungkin saja membuat susunan pengukuran baru yang dapat merepresentasikan setiap kelompok dari variable-variable tersebut.

Kita perkenal analisis factor sebagai teknik multivariate yang pertama sebab analisis dapat memainkan peranan penting dalam penerapan teknik multivariate yang lain. Dengan perkataan lain, analisis factor menyediakan alat menggambarkan secara lebih mantap struktur inter-keterhubungan (Korelasi) antara variable dengan jumlah yang sangat banyak dengan mendefinisikan kumpulan variable yang saling berkorelasi tinggi yang selanjutnya disebut **Faktor**. Grup-group dari variable-variable tersebut (factor-faktor), yang didefinisikan saling berkorelasi dengan korelasi yang tinggi, diasumsikan mewakili dimensi antar data.

II. LANDASAN TEORI

A. Model Orthogonal Faktor

Vektor acak \mathbf{X} teramati, dengan p komponen, mempunyai parameter-parameter m dan \mathbf{S} masing-masing merupakan vektor rata-rata dan matriks kovariansi. Dalil yang menyatakan bahwa model factor \mathbf{X} adalah terjadinya

independen linear sekumpulan kecil variable acak tak teramati, F_1, F_2, L, F_m , disebut common factor, dan p sumber variasi tambahan e_1, e_2, L, e_m disebut error atau kadang disebut juga factor khusus. Model analisis factor adalah:

$$\begin{matrix} X_1 - m_1 = 1_{11}F_1 + 1_{12}F_2 + L + 1_{1m}F_m + e_1 \\ X_2 - m_2 = 1_{21}F_1 + 1_{22}F_2 + L + 1_{2m}F_m + e_2 \\ \vdots \\ X_p - m_p = 1_{p1}F_1 + 1_{p2}F_2 + L + 1_{pm}F_m + e_p \end{matrix} \quad (0.1)$$

yang dapat dituliskan dalam notasi matriks, yaitu

$$\mathbf{X} - \mathbf{m} = \mathbf{L} \mathbf{F} + \mathbf{e}$$

$\begin{matrix} (p \times 1) & & (p \times m) & (m \times 1) & & (p \times 1) \end{matrix}$

Koefisien 1_{ij} disebut loading dari variable ke- i factor ke- j , dengan demikian matriks \mathbf{L} merupakan matriks factor loading. Perhatikan bahwa factor khusus ke- i , e_i , hanya berkaitan dengan respon ke- i dari X_i .

Jumlah besaran yang tak teramati yang cukup banyak, verifikasi secara langsung model factor dari pengamatan X_1, X_2, L, X_p kurang membantu. Namun demikian dengan beberapa tambahan asumsi seputar vektor acak \mathbf{F} dan \mathbf{e} , model **Error! Reference source not found.** memberikan implikasi hubungan kovariansi tertentu, yang dapat diperiksa. Asumsikan bahwa

$$E(\mathbf{F}) = \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \begin{matrix} (m \times 1) \end{matrix}, \text{cov}(\mathbf{F}) = E(\mathbf{F}\mathbf{F}^T) = \mathbf{I} \begin{matrix} (m \times m) \end{matrix}, E(\mathbf{e}) = \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \begin{matrix} (p \times 1) \end{matrix}, \text{cov}(\mathbf{e}) = E(\mathbf{e}\mathbf{e}^T) = \mathbf{M} = \begin{matrix} M_1 & 0 & L & 0 \\ 0 & M_2 & L & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & L & M_p \end{matrix} \begin{matrix} (p \times p) \end{matrix} \quad (0.2)$$

dan bahwa \mathbf{F} dan \mathbf{e} adalah independen, sehingga

$$\text{cov}(\mathbf{e}, \mathbf{F}) = E(\mathbf{e}\mathbf{F}^T) = \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \begin{matrix} (p \times m) \end{matrix}$$

Asumsi-asumsi tersebut dan persamaan

Error! Reference source not found. disebut model factor orthogonal.

Struktur kovariansi untuk model factor orthogonal

$$1. \text{cov}(\mathbf{X}) = \mathbf{L}\mathbf{L}^T + \mathbf{M}$$

atau

$$\begin{matrix} \text{var}(X_i) = 1_{i1}^2 + L + 1_{im}^2 + M_i \\ \text{cov}(X_i, X_j) = 1_{i1}1_{j1} + L + 1_{im}1_{jm} \end{matrix} \quad (0.3)$$

$$2. \text{cov}(\mathbf{X}, \mathbf{F}) = \mathbf{L}$$

atau

$$\text{cov}(X_i, F_j) = 1_{ij} \quad (0.4)$$

Bagian dari variansi variable ke- i yang berkontribusi dengan m common factor disebut **communality**. Bagian dari $\text{var}(X_i) = s_{ii}$ tersebut yang menyatakan factor tertentu sering disebut uniqueness atau variansi tertentu.

B. Metode Estimasi

1. Metode principal component

Analisis principal component factor dari matriks kovariansi sampel S adalah ditetapkan dalam model pasangan nilai eigen-vektor eigen $(\hat{l}_1, \hat{e}_1), (\hat{l}_2, \hat{e}_2), \dots, (\hat{l}_p, \hat{e}_p)$ dimana $\hat{l}_1 \geq \hat{l}_2 \geq \dots \geq \hat{l}_p$. Misalkan $m < p$ adalah jumlah common factor. maka matriks dari estimasi loading factor $\{\varphi_{ij}\}$ diberikan oleh

$$\hat{F} = \frac{1}{\sqrt{\hat{l}_1}} \hat{e}_1 \sqrt{\hat{l}_2} \hat{e}_2 \dots \sqrt{\hat{l}_m} \hat{e}_m \quad (0.5)$$

Estimasi variansi tertentu diberikan oleh elemen diagonal dari matriks $S - \hat{F}\hat{F}'$, sehingga

$$\hat{M} = \begin{pmatrix} M_1 & 0 & L & 0 \\ 0 & M_2 & L & 0 \\ M & M & O & M \\ 0 & 0 & L & M_p \end{pmatrix} \text{ dengan } \hat{M}_i = s_{ii} - \sum_{j=1}^m \varphi_{ij}^2 \quad (0.6)$$

Communalities diestimasi sebagai

$$\hat{\rho}_i^2 = \varphi_{i1}^2 + \varphi_{i2}^2 + L + \varphi_{im}^2 \quad (0.7)$$

Analisis factor principal component dari matriks korelasi sampel adalah diperoleh dengan terlebih dahulu menghitung R .

2. Maksimum Likelihood

Jika factor umum F dan factor tertentu ε dapat diasumsikan berdistribusi normal, maka estimasi dengan maksimum likelihood dari factor loading dan variansi tertentu dimungkinkan untuk didapatkan. Pada saat F_j dan e_j merupakan gabungan normal, pengamatan $X_j - m = LF_j + e$ juga normal, fungsi like likelihood

$$L(m, S) = (2p)^{-\frac{np}{2}} |S|^{-\frac{n}{2}} e^{-\frac{1}{2} \sum_{j=1}^p \frac{(x_j - \bar{x})(x_j - \bar{x})}{\sigma_j^2} + n(\bar{x} - m)(\bar{x} - m)} \quad (0.8)$$

bergantung pada L dan Y melalui $S = LL^T + Y$

C. Rotasi Faktor

Suatu transformasi factor loading disebut factor rotasi. Jika \hat{L} merupakan matriks estimasi factor loading berukuran $p' \times m$ yang diperoleh dari metode ekstraksi factor (pca, atau yang lainnya), maka

$$\hat{L}^* = \hat{L}T \quad \text{dimana} \quad TT^T = T^T T = 1 \quad (0.9)$$

adalah matriks berukuran $p' \times m$ dari loading yang dirotasikan.

Oleh karena loading asal agak sulit untuk diinterpretasikan, biasanya untuk memudahkan membacanya, factor loading tersebut dirotasikan hingga mendapat struktur yang sederhana.

D. Skor Faktor

Nilai estimasi dari common factor disebut factor score. Besaran ini biasanya digunakan untuk tujuan diagnosis. Untuk mendapatkan skor factor disini akan menggunakan metode WLS, yaitu: untuk vektor rata-rata μ , loading factor L dan variansi Y diketahui untuk factor model

$$\underset{(p' \times 1)}{X} - \underset{(p' \times 1)}{m} = \underset{(p' \times 1)(p' \times 1)}{L} \underset{(p' \times 1)}{F} + \underset{(p' \times 1)}{e}$$

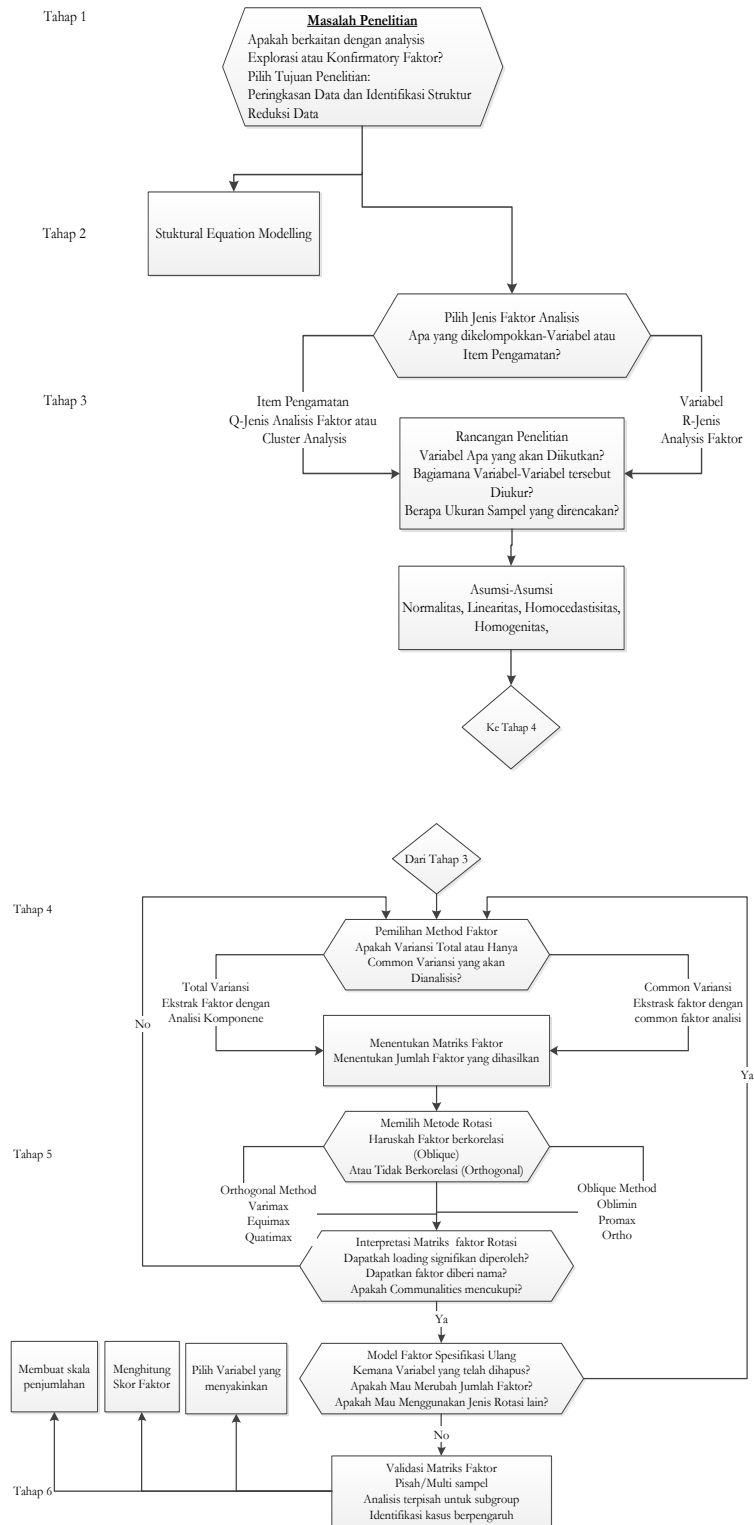
Jumlah kuadrat error, terboboti adalah

$$\sum_{i=1}^p \frac{e_i^2}{y_i} = e^T Y^{-1} e = (x - m - Lf)^T Y^{-1} (x - m - Lf) \quad (0.10)$$

dengan menggunakan teknik yang dikemukakan oleh Bartlett, estimasi \hat{f} dari f untuk meminimumkan (0.10), diperoleh

$$\hat{f} = (LYL)^{-1} LY^{-1} (x - m)$$

III. PROSES KEPUTUSAN ANALISIS FAKTOR



IV. ANALISIS FAKTOR MENGGUNAKAN SOFTWARE MINITAB

Diberikan data berikut

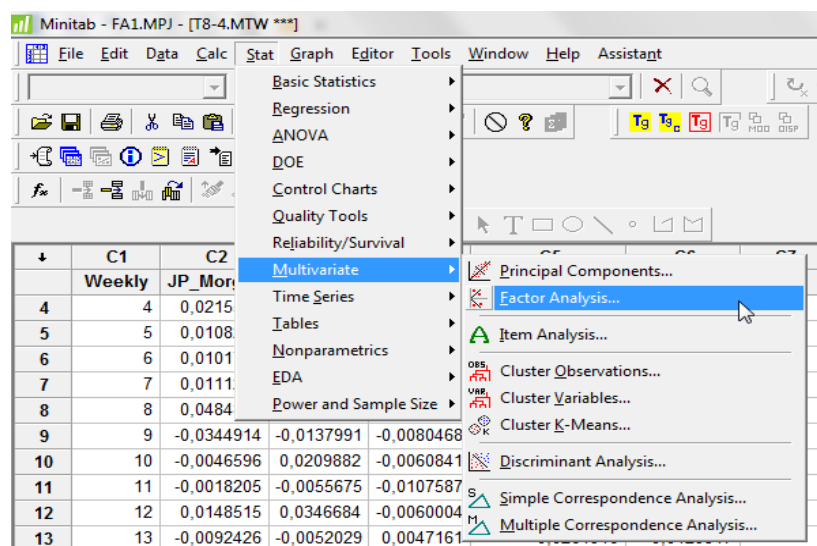
Pop	School	Employ	Health	Home
5,935	14,2	2,265	2,27	2,91
1,523	13,1	0,597	0,75	2,62
2,599	12,7	1,237	1,11	1,72
4,009	15,2	1,649	0,81	3,02
4,687	14,7	2,312	2,50	2,22
8,044	15,6	3,641	4,51	2,36
2,766	13,3	1,244	1,03	1,97
6,538	17,0	2,618	2,39	1,85
6,451	12,9	3,147	5,52	2,01
3,314	12,2	1,606	2,18	1,82
3,777	13,0	2,119	2,83	1,80
1,530	13,8	0,798	0,84	4,25
2,768	13,6	1,336	1,75	2,64
6,585	14,9	2,763	1,91	3,17

A. Langkah-Langkah Analisis factor

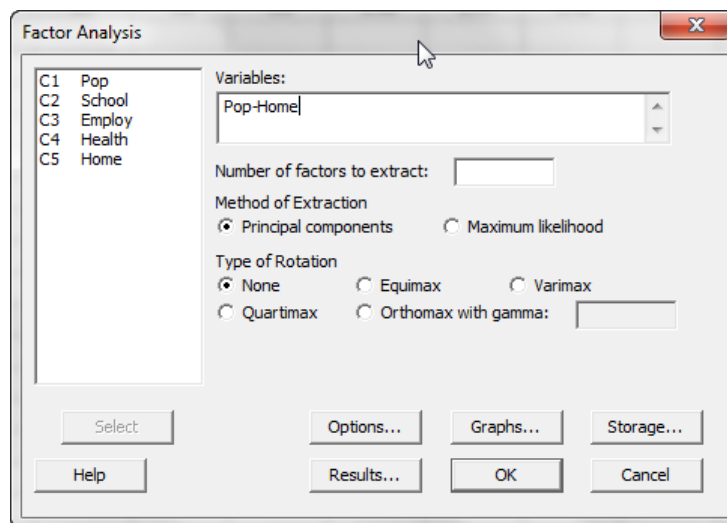
1. Menggunakan Metode maksimum Likelihood

a. Langkah Analisis dengan Minitab

Pilihan Stat > Multivariate > Factors Analysis



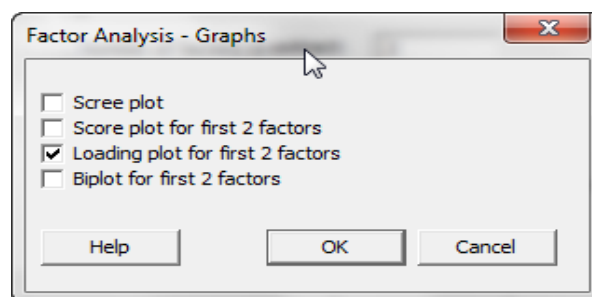
Selanjutnya akan tampil jendela Factor Analysis..



Pada jendela Factor Analysis:

- a) Pilih Variabel **Pop, School, Employ, Health, dan Home**, pindah ke dalam kotak **Variables**
- b) Pada bagian **Number of Factors to extract** ketik **2**
- c) Di bawah **method of extraction**, Pilih **Maximum likelihood**
- d) Di bawah **Type of rotation**, pilih **Varimax**
- e) Klik tombol **Graphs** lalu pilih **Loading plot for first 2 factor**, lalu klik tombol **OK**.

Lihat gambar berikut:



- f) Klik Tombol **Result**, lalu pilih short loading, lalu klik **OK**.

Lihat gambar berikut

- g) Klik **OK**. Untuk memulai proses

b. Hasilnya

Factor Analysis: Pop; School; Employ; Health; Home

Principal Component Factor Analysis of the Correlation Matrix

Unrotated Factor Loadings and Communalities

Variable	Factor1	Factor2	Communality
Pop	0,972	0,149	0,967
School	0,545	0,715	0,808
Employ	0,989	0,005	0,978
Health	0,847	-0,352	0,841
Home	-0,303	0,797	0,726
Variance	3,0289	1,2911	4,3200
% Var	0,606	0,258	0,864

Rotated Factor Loadings and Communalities

Varimax Rotation

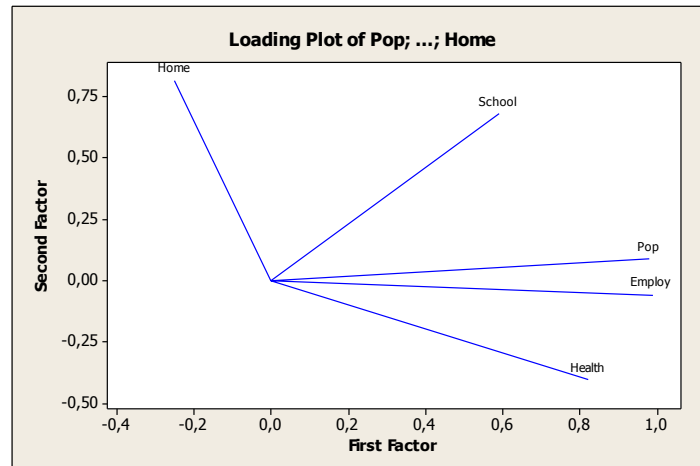
Variable	Factor1	Factor2	Communality
Pop	0,979	0,085	0,967
School	0,591	0,677	0,808
Employ	0,987	-0,060	0,978
Health	0,822	-0,406	0,841
Home	-0,251	0,815	0,726
Variance	3,0215	1,2986	4,3200
% Var	0,604	0,260	0,864

Sorted Rotated Factor Loadings and Communalities

Variable	Factor1	Factor2	Communality
Employ	0,987	-0,060	0,978
Pop	0,979	0,085	0,967
Health	0,822	-0,406	0,841
Home	-0,251	0,815	0,726
School	0,591	0,677	0,808
Variance	3,0215	1,2986	4,3200
% Var	0,604	0,260	0,864

Factor Score Coefficients

Variable	Factor1	Factor2
Pop	0,328	0,094
School	0,216	0,541
Employ	0,326	-0,018
Health	0,261	-0,290
Home	-0,060	0,622



c. Interpretasi

1) Hasil Perhitung menampilkan 3 tabel Loading dan Communalities

a) Unrotated

Factor Unrotated menjelaskan variabilitas data sebesar 79% (Lihat dibawah communalities).

Nilai communalities mengindikasikan bahwa semua variable kecuali Home dapat digambarkan dengan baik oleh kedua factor (Communalities 0.202 untuk home, 0.875 untuk variable yang lain).

b) Rotated

Total variabilitas yang digambarkan oleh factor tidak berubah dengan rotasi, akan nanti setelah rotasi dilakukan, factor-faktor tersebut nampak cenderung sama dalam variabilitas.

c) Sorted and rotated

2) Interpretasi factor,

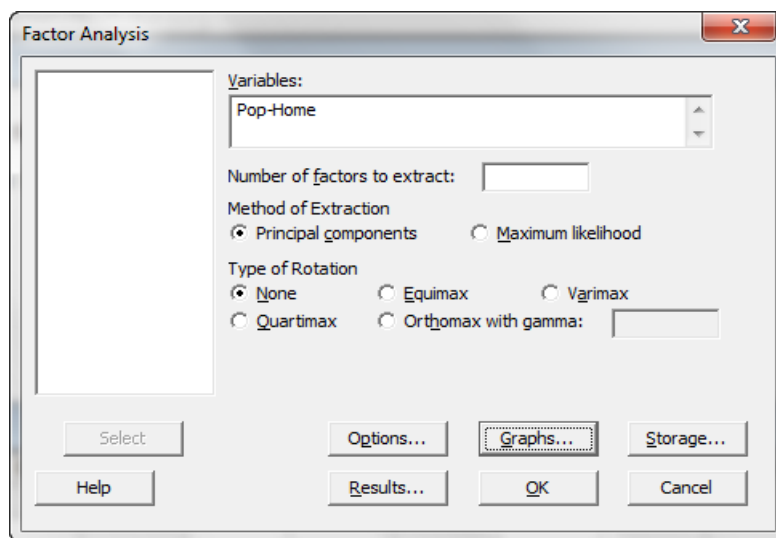
Factor pertama loading positif pada ukuran populasi dan pada dua variable, employ dan health.

Negatif load pada nilai home, tapi mungkin memberikan pengaruh besar oleh satu titik. Sehingga dapat dipertimbangkan factor 1 menjadi health care – factor ukuran populasi.

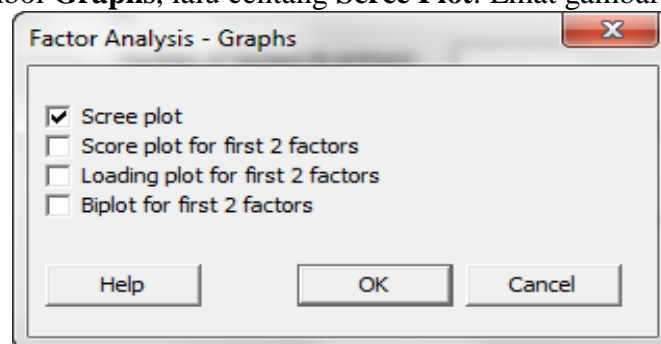
Factor kedua, dapat mempertimbangkan Education-ukuran populasi sebagai factor.

2. Menggunakan Metode Principal Component

Dengan Data yang sama dari table. Pada langkah ke-3 dari latihan sebelumnya, Pada jendela Factor Analysis,



- Dibawah **Method of extraction**, pilih **Principal Component**
- Klik tombol **Graphs**, lalu centang **Scree Plot**. Lihat gambar berikut



- Klik OK
- Klik OK

a. Hasilnya

Factor Analysis: Pop; School; Employ; Health; Home

Principal Component Factor Analysis of the Correlation Matrix

Unrotated Factor Loadings and Communalities

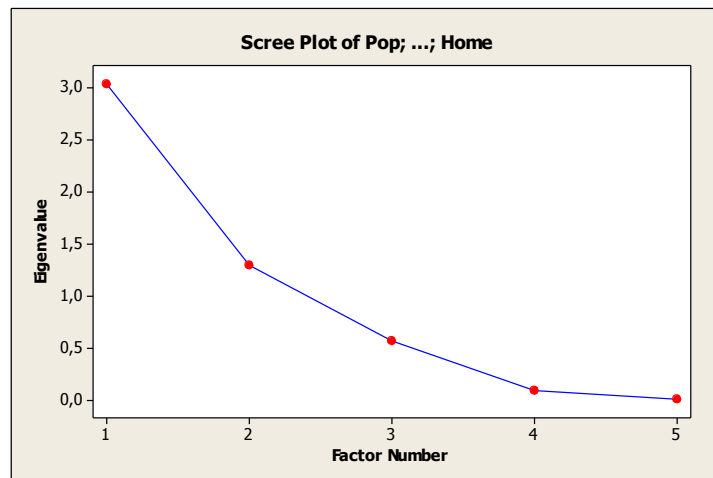
Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Communality
Pop	0,972	0,149	-0,006	-0,170	0,067	1,000
School	0,545	0,715	0,415	0,140	-0,001	1,000
Employ	0,989	0,005	-0,089	-0,083	-0,085	1,000
Health	0,847	-0,352	-0,344	0,200	0,022	1,000
Home	-0,303	0,797	-0,523	-0,005	-0,002	1,000
Variance	3,0289	1,2911	0,5725	0,0954	0,0121	5,0000
% Var	0,606	0,258	0,114	0,019	0,002	1,000

Sorted Unrotated Factor Loadings and Communalities

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Communality
Employ	0,989	0,005	-0,089	-0,083	-0,085	1,000
Pop	0,972	0,149	-0,006	-0,170	0,067	1,000
Health	0,847	-0,352	-0,344	0,200	0,022	1,000
Home	-0,303	0,797	-0,523	-0,005	-0,002	1,000
School	0,545	0,715	0,415	0,140	-0,001	1,000
Variance	3,0289	1,2911	0,5725	0,0954	0,0121	5,0000
%Var	0,606	0,258	0,114	0,019	0,002	1,000

Factor Score Coefficients

Variable	Factor1	Factor2	Factor3	Factor4	Factor5
Pop	0,321	0,116	-0,011	-1,782	5,511
School	0,180	0,553	0,726	1,466	-0,060
Employ	0,327	0,004	-0,155	-0,868	-6,988
Health	0,280	-0,272	-0,601	2,098	1,829
Home	-0,100	0,617	-0,914	-0,049	-0,129



b. Interpretasi

Dari grafik scree plot, terdapat lima factor yang mungkin. Akan tetapi perhatikan nilai eigen untuk factor kelima, bahwa % variasi dari factor kelima dan keempat berkisar 0,002 dan 0,019 sehingga kedua dapat dihapuskan karena tidak cukup memadai. Factor kesatu dan kedua, % variasi adalah 86% dan untuk tiga factor pertama % var sebesar 98%.

Pertanyaanya: berapa factor yang akan diambil? Dua atau tiga? Untuk menjawab pertanyaan tersebut, kita lakukan analisis lebih lanjut dengan menggunakan metode analisis Kluster pengamatan.

DAFTAR PUSTAKA

- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2012). *Multivariate Data Analysis* (Seventh Edition ed.). New Jersey: John Wiley & Sons.
- Harris, R. J. (2001). *A Primer of Multivariate Statistics* (Third Edition ed.). New Jersey: Lawrence Erlbaum Associates.
- Rencher, A. C., & Christensen, W. F. (2012). *Method of Multivariate Analysis* (Third Edition ed.). New Jersey: John Wiley & Sons Inc.
- Seber, G. A. (2007). *A Matrix Handbook for Statisticians*. New York: A John Wiley & Sons. Inc.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (Sixth Edition ed.). New Jersey: Pearson.

Timm, N. H. (2002). *Applied Multivariate Analysis*. New York: Springer.

Wichern, D. W., & Johnson, R. A. (2007). *Applied Multivariate Statistical Analysis* (Sixth Edition ed.). New Jersey: Pearson Prentice Hall.