

Exploratory factor analyses (Phân tích nhân tố khám phá)

Nguyễn Trương Nam
Nguyễn Hồng Phương

Copyright – Bản quyền thuộc về tác giả và thongke.info. Khi sử dụng một phần hoặc toàn bộ bài giảng đề nghị mọi người trích dẫn: tên tác giả và thongke.info. Ví dụ: Nguyễn Thị Linh – Thongke.info.

Mục tiêu

- Giới thiệu về factor analysis
- Các bước trong factor analysis
- Thao tác thực hiện
- Đọc kết quả
- Các phân tích khác sau khi thực hiện factor analysis

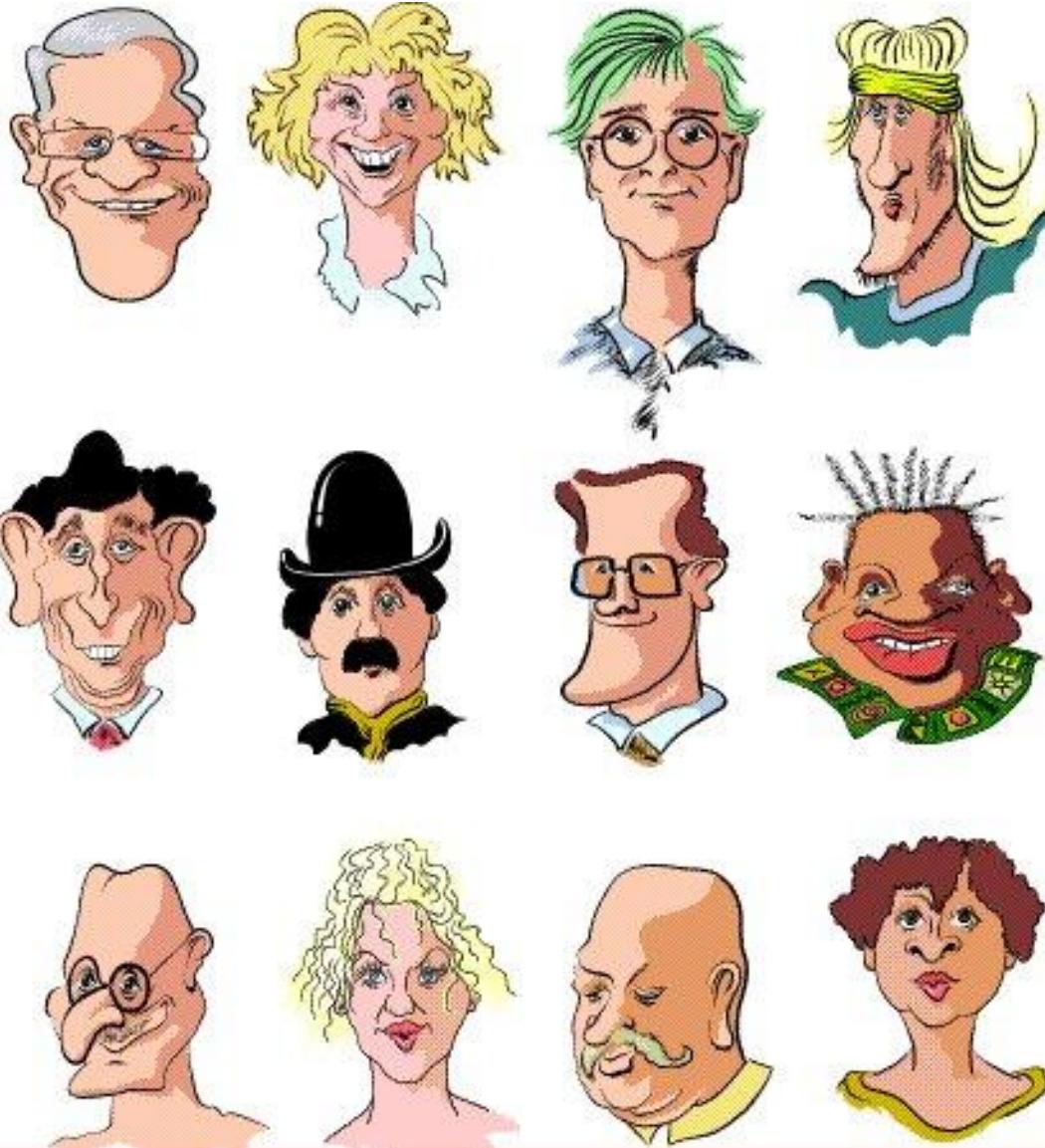
Tại sao cần Factor analysis

- Một số biến có thể đo trực tiếp: tốc độ, chiều cao, cân nặng
- Một số biến khác không thể đo trực tiếp bằng một câu hỏi: tính sáng tạo, hạnh phúc, sự hài lòng, sự thoái mái....
- Các biến này cần được tạo ra một cách gián tiếp từ nhiều câu hỏi khác nhau
- Factor analysis sẽ giúp tạo ra các biến này

Factor analysis là gì?

- FA là các phương pháp rút gọn data
 - Tìm mối liên quan của các **biến liên tục** dựa trên mối liên quan của chúng
 - Phân tích nhiều biến và giải thích chúng bằng vài nhân tố hoặc thành tố
 - Các biến có liên quan với nhau được nhóm lại với nhau và tách ra khỏi các biến khác mà nó ít liên quan

Bạn nhóm các gương mặt này theo các cách nào?



Factor Analysis

| Biến quan sát | Nhân tố | | |
|--------------------------------------|---------|-------------|--------|
| | Dịch vụ | Sự tiện lợi | Giá cả |
| Thái độ của bác sĩ , y tá niềm nở | 0.59 | | |
| Mở cửa đúng giờ | 0.65 | | |
| Giúp bệnh nhân qua điện thoại | 0.62 | | |
| Khoảng cách từ nhà đến bệnh viện gần | | 0.65 | |
| Có chỗ đậu xe | | 0.86 | |
| Có cửa hàng thuốc bệnh cạnh | | 0.83 | |
| Có dịch vụ buổi tối và cuối tuần | | 0.72 | |
| Không phải trả tiền khám cho trẻ em | | | 0.73 |
| Chi phí nằm nội trú thấp | | | 0.85 |
| Được giảm giá cho một số xét nghiệm | | | 0.91 |

Điều kiện

- Các biến phải có liên quan với nhau
- Nếu mối liên quan mà nhỏ- không thích hợp cho phương pháp Factor Analysis
- Test ma trận các mối liên quan bằng:
 - KMO test
 - Bartlett's test of sphericity

Các bước trong Factor Analysis

Factor analysis thường có 4 bước:

- Bước 1: Tính ma trận các mối liên quan cho tất cả các biến (correlation matrix)
- Bước 2 : Factor extraction
- Bước 3 : Factor rotation
- Bước 4 : Ra quyết định cuối cùng về số nhân tố cần giữ lại

Các bước trong Factor Analysis: The Correlation Matrix

- **Bước 1: the correlation matrix**

- Tạo correlation matrix cho tất cả các biến
- Xác định các biến mà không có liên quan với biến khác
- Nếu mối liên quan giữa các biến nhỏ, chúng có thể không chung một nhân tố- FA không phù hợp
- Correlation coefficients ≥ 0.3 - OK.

Các bước trong Factor Analysis: The Correlation Matrix

- **The Kaiser-Meyer-Olkin (KMO)**
- **Bartlett Test of Sphericity:**

KMO and Bartlett's Test

| | | |
|--|--------------------|---------|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .801 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 632.486 |
| | df | 21 |
| | Sig. | .000 |

- KMO là một chỉ tiêu để xem xét sự thích hợp của FA Nếu $0.5 \leq KMO \leq 1$ thì phân tích nhân tố là thích hợp
- Kiểm định Bartlett xem xét giả thuyết H0: Độ tương quan giữa các biến số quan sát bằng 0. Nếu kiểm định này có ý nghĩa thống kê ($sig < 0.05$) thì các biến có tương quan với nhau trong tổng thể

Các bước trong Factor Analysis: Factor Extraction

- **Bước 2: Factor extraction**

- Mục đích chính của bước này là xác định các nhân tố.
- Thường dùng phương pháp **Principal components analysis** để xác định các nhân tố.
 - The 1st principal component là phức hợp giải thích được nhiều biến thiên nhất trong quần thể (1st extracted factor), sau đó giảm dần ở nhân tố thứ 2, 3...

Các bước trong Factor Analysis: Factor Extraction

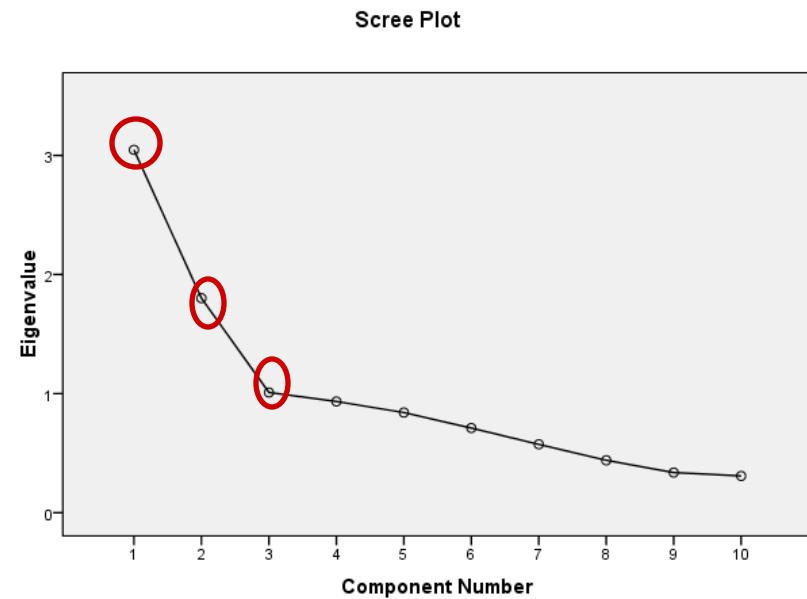
- Để xác định giữ lại bao nhiêu factor, căn cứ vào 2 yếu tố:
 - Eigen Values >1
 - The Scree Plot.

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | |
|-----------|---------------------|---------------|--------------|-------------------------------------|---------------|--------------|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| | 3.046 | 30.465 | 30.465 | 3.046 | 30.465 | 30.465 |
| 1 | 3.046 | 30.465 | 30.465 | 3.046 | 30.465 | 30.465 |
| 2 | 1.801 | 18.011 | 48.476 | 1.801 | 18.011 | 48.476 |
| 3 | 1.009 | 10.091 | 58.566 | 1.009 | 10.091 | 58.566 |
| 4 | .934 | 9.336 | 67.902 | | | |
| 5 | .840 | 8.404 | 76.307 | | | |
| 6 | .711 | 7.107 | 83.414 | | | |
| 7 | .574 | 5.737 | 89.151 | | | |
| 8 | .440 | 4.396 | 93.547 | | | |
| 9 | .337 | 3.368 | 96.915 | | | |
| 10 | .308 | 3.085 | 100.000 | | | |

Extraction Method: Principal Component Analysis.

Các bước trong Factor Analysis: Factor Extraction- scree plot

- Độ dốc cho thấy các factor lớn.
- Khi hết độ dốc, thường các factor còn lại có giá trị Eigen <1 .
- Ngoài các test thống kê, cần dựa vào thực tiễn và y văn
- Giai đoạn này chưa kết luận số nhân tố.



Steps in Factor Analysis: Factor Rotation

- **Bước 3: Factor rotation.**
- Ở bước này các factors được rotated để nó có ý nghĩa hơn
- Có nhiều phương pháp rotation
- Phương pháp hay dùng nhất là **Varimax rotations**.

Steps in Factor Analysis: Factor Rotation

Chọn các factor có trị số lớn nhất và nhóm chúng thành nhóm, rồi đặt tên cho chúng

Rotated Component Matrix^a

| | Component | | |
|--|-----------|-------|-------|
| | 1 | 2 | 3 |
| I discussed my frustrations and feelings with person(s) in school | .803 | .186 | .050 |
| I tried to develop a step-by-step plan of action to remedy the problems | .270 | .304 | .694 |
| I expressed my emotions to my family and close friends | .706 | -.036 | .059 |
| I read, attended workshops, or sought some other educational approach to correct the problem | .050 | .633 | .145 |
| I tried to be emotionally honest with myself about the problems | .042 | .685 | .222 |
| I sought advice from others on how I should solve the problems | .792 | .117 | -.038 |
| I explored the emotions caused by the problems | .248 | .782 | -.037 |
| I took direct action to try to correct the problems | -.120 | -.023 | .772 |
| I told someone I could trust about how I felt about the problems | .815 | .172 | -.040 |
| I put aside other activities so that I could work to solve the problems | -.014 | .155 | .657 |

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Steps in Factor Analysis: Quyết định cuối cùng

• Bước 4: Quyết định

- Chọn các factors có ý nghĩa thực tiễn nhất
- Chọn các nhóm biến số có chỉ số lớn cho cùng một factor
- Đặt tên cho factor theo ý nghĩa của các biến

Sau đó:

- Các factor này có ý nghĩa không? Có phù hợp với y văn trước đây không?
- Nên dùng chỉ số factor hay chỉ số thường trong các phân tích tiếp theo?

“Một Factor tốt”

- Makes sense
- will be easy to interpret
- simple structure
- Lacks complex loadings

Thực hành trên máy

Bước 1: Từ thanh menu của SPSS chọn
Analyze

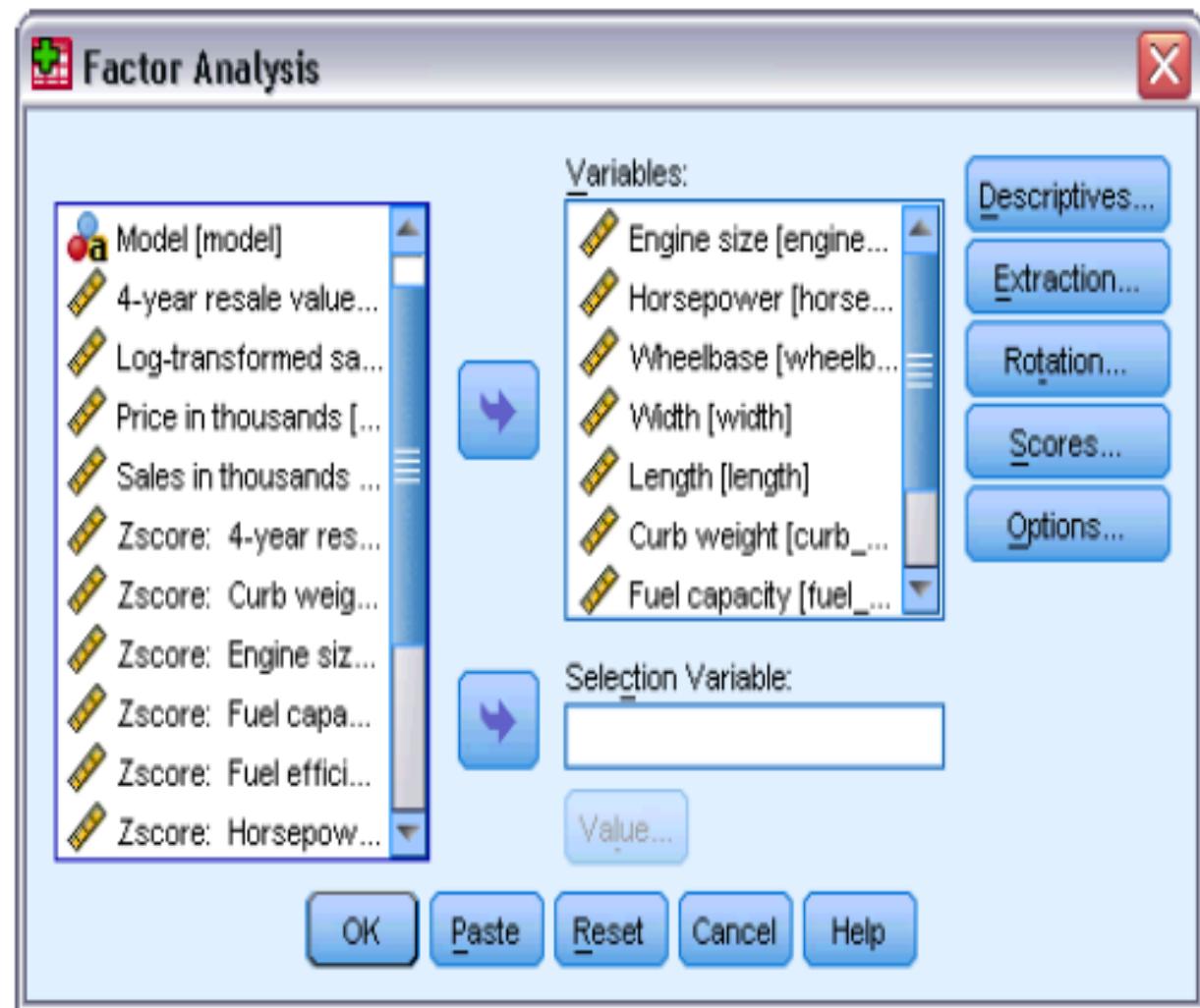
Dimention reduction
Factor

Khi hộp Factor analyses xuất hiện:

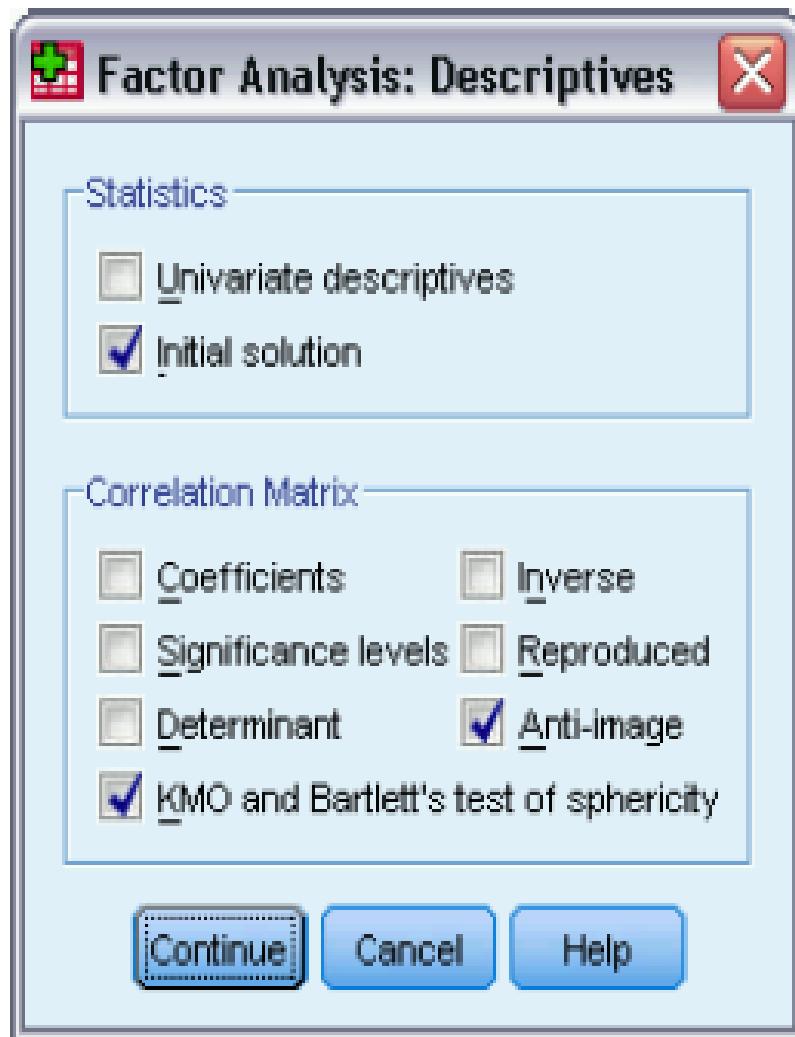
- Đưa các biến cần phân tích thành nhân tố vào khung

Variables

- Sau đó lần lượt chọn các nút:
- Descriptives
- Extraction
- Rotation
- Scores



Factor Analysis Descriptive

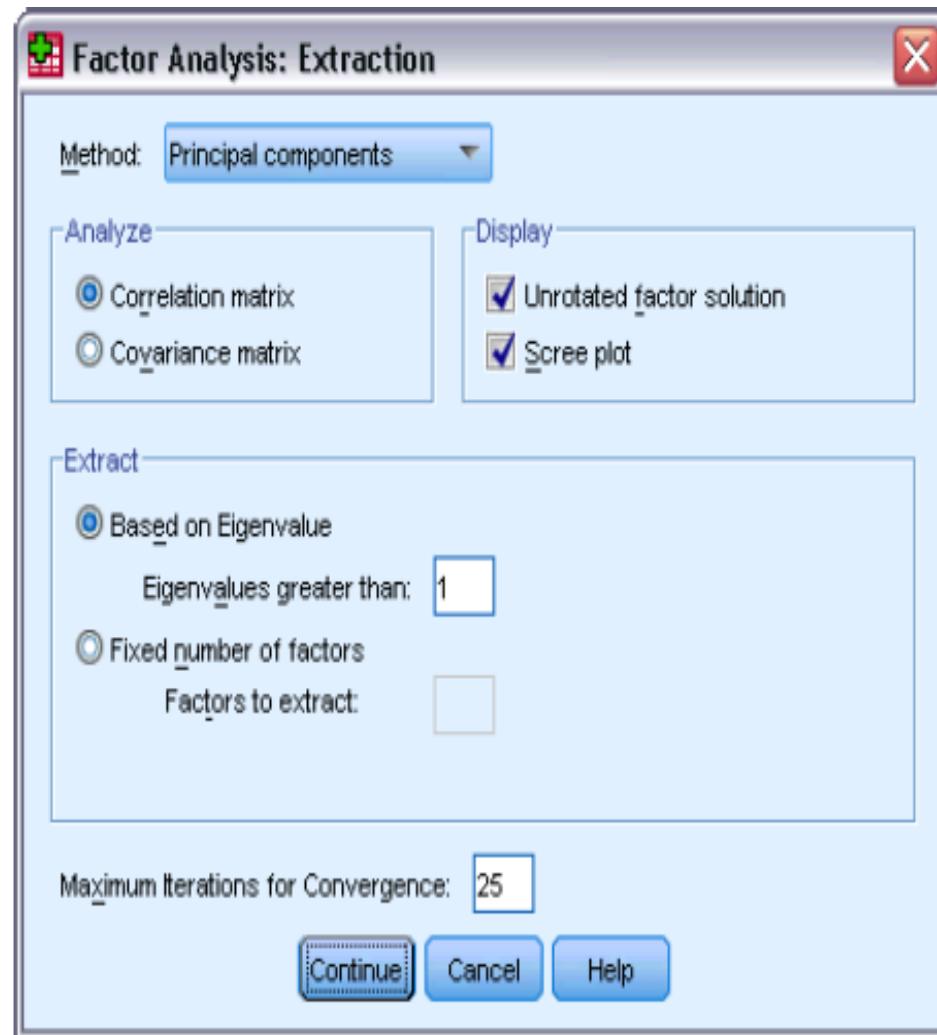


Statistics:

- Univariate descriptive: mean, SD
- Initial solution: communalities, eigenvalues, % of variance explained

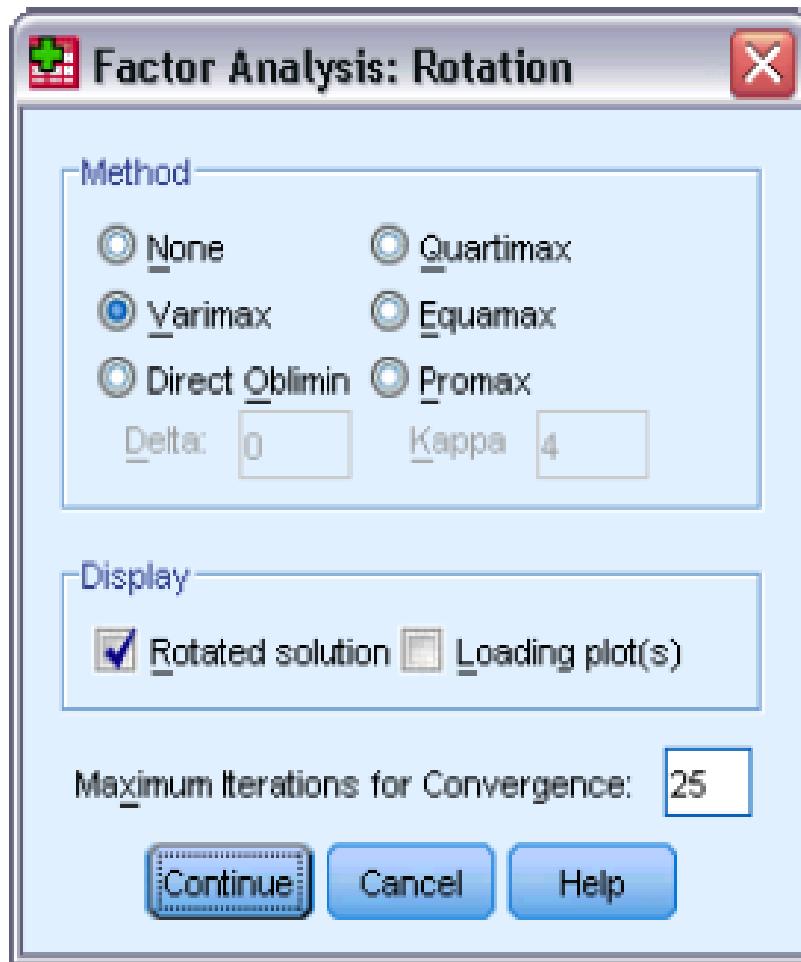
KMO and Bartlett's Test of Sphericity: Test factor model is appropriate

Factor Analysis Extraction



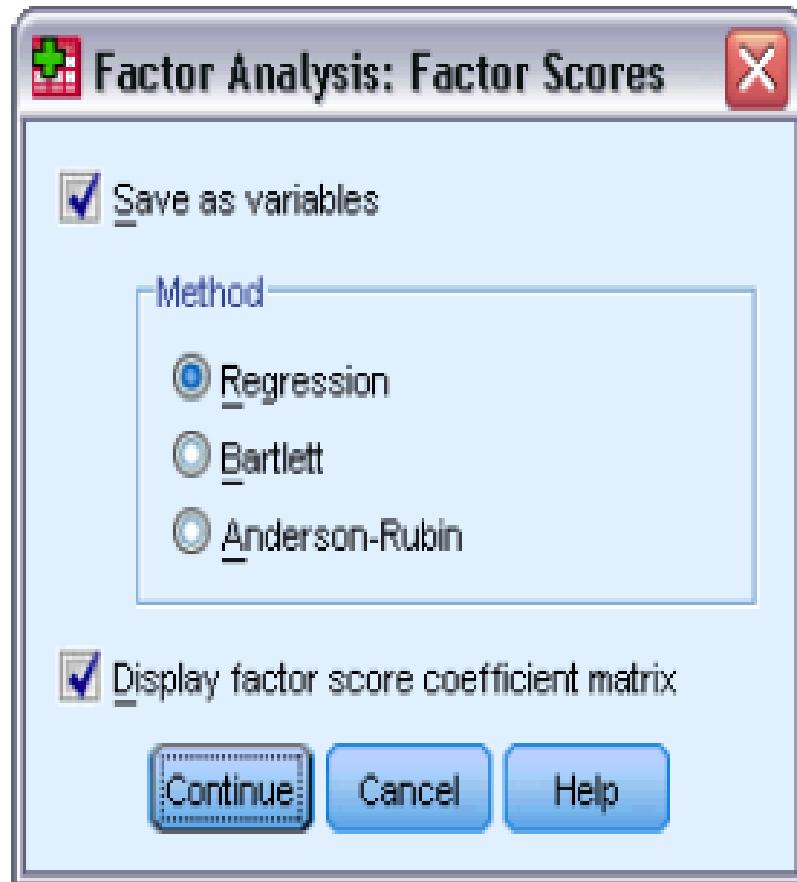
Khi chọn **Extraction**, chọn phương pháp mặc định của SPSS là **Principle components**

Factor Analysis Rotation



Trong hộp Rotation, đánh dấu **Varimax** nếu phương pháp là **Principle components**

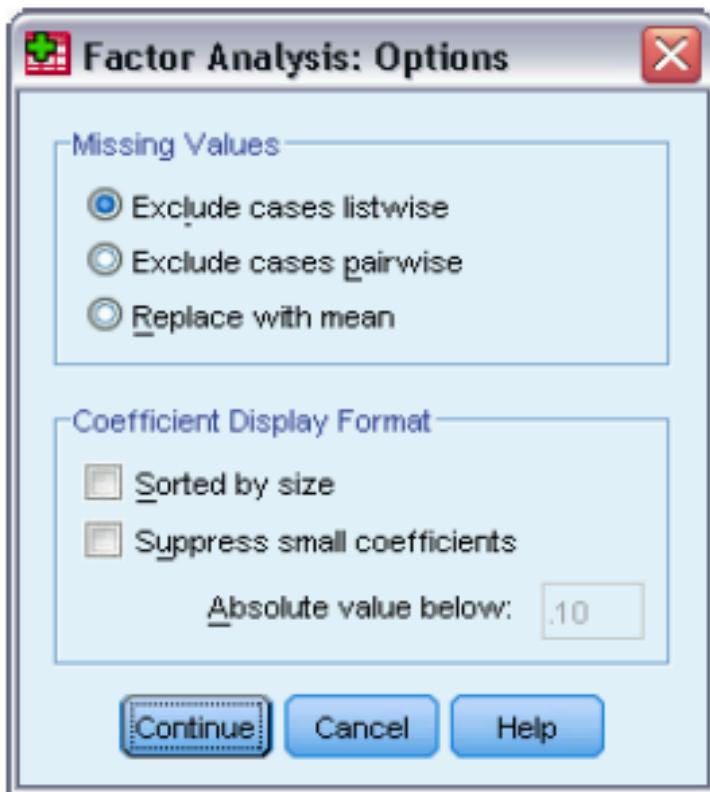
Factor Analysis Scores



Khi bấm nút **Scores**, hộp **Factor Analyes: Factor Scores** xuất hiện

Chọn **Save as variables** nếu bạn muốn lưu lại nhân số (đã chuẩn hóa) của từng nhân tố

Factor Analysis Options



- Chọn **Sorted by size** để sắp xếp các biến quan sát trong cùng một nhân tố đứng gần nhau
- Chọn **Suppress absolute value less than** nếu không thể hiện các nhân tố có giá trị tuyệt đối nhỏ hơn một giá trị nào đó (ví dụ 0.3)

Đọc kết quả

KMO and Bartlett's Test

| | |
|--|--------------------|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | .807 |
| Bartlett's Test of Sphericity | Approx. Chi-Square |
| | df |
| | Sig. |

- KMO $0.5 \leq KMO \leq 1$
- Bartlett test of Sphericity < 0.05 : các biến có tương quan với nhau trong tổng thể

Đọc kết quả

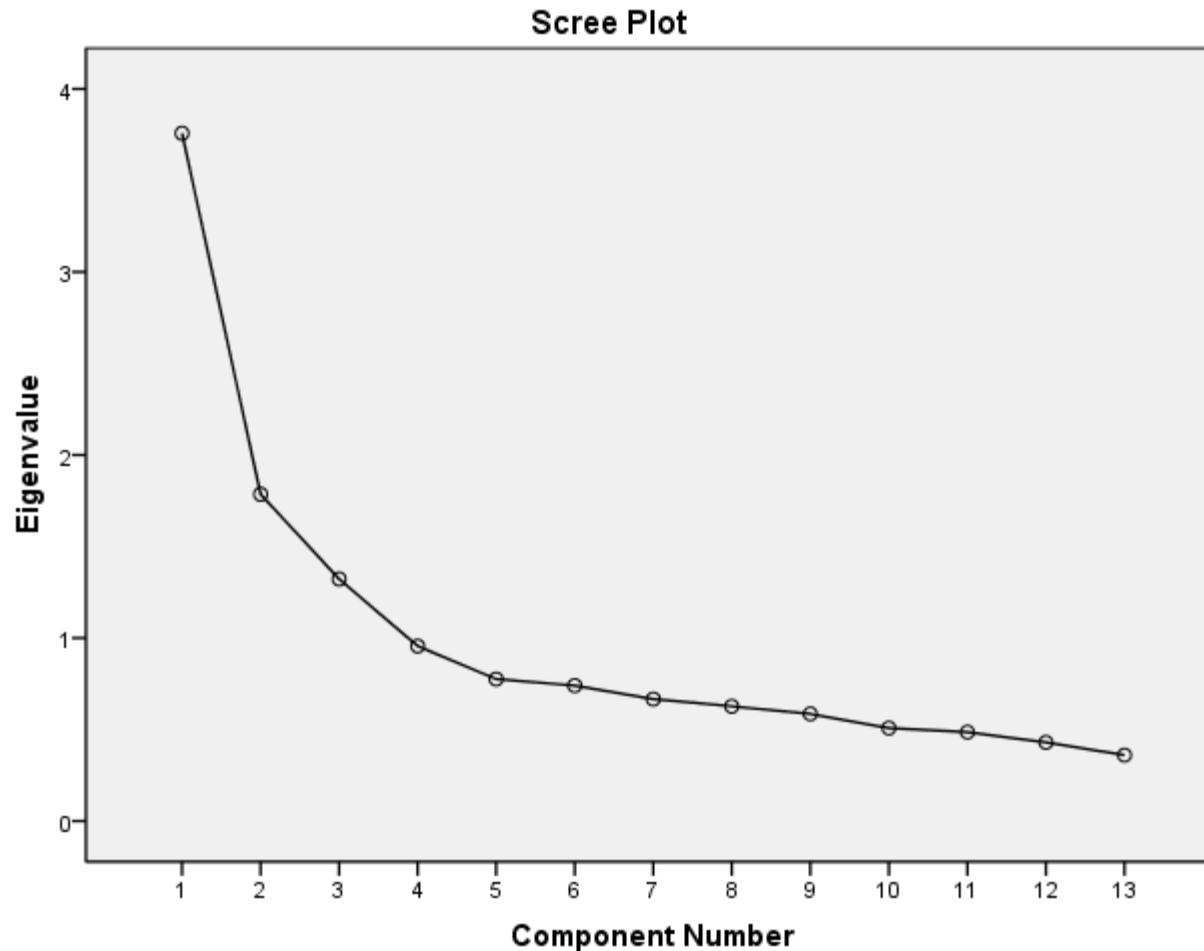
Total Variance Explained

| Component | Initial Eigenvalues | | | Rotation Sums of Squared Loadings | | |
|-----------|---------------------|---------------|--------------|-----------------------------------|---------------|--------------|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 3.758 | 28.905 | 28.905 | 2.561 | 19.699 | 19.699 |
| 2 | 1.785 | 13.734 | 42.639 | 2.184 | 16.800 | 36.499 |
| 3 | 1.322 | 10.172 | 52.811 | 2.121 | 16.312 | 52.811 |
| 4 | .956 | 7.354 | 60.165 | | | |
| 5 | .776 | 5.966 | 66.131 | | | |
| 6 | .740 | 5.694 | 71.825 | | | |
| 7 | .667 | 5.127 | 76.952 | | | |
| 8 | .627 | 4.823 | 81.776 | | | |
| 9 | .585 | 4.502 | 86.278 | | | |
| 10 | .508 | 3.908 | 90.186 | | | |
| 11 | .486 | 3.736 | 93.923 | | | |
| 12 | .430 | 3.306 | 97.229 | | | |
| 13 | .360 | 2.771 | 100.000 | | | |

Extraction Method: Principal Component Analysis.

Con số 53% cho biết 3 nhân tố giải thích được 53% biến thiên của các biến quan sát hay của dữ liệu

Đọc kết quả



Đọc kết quả

Factor loading lớn nhất phải lớn hơn hoặc bằng 0.5. Nếu nó nhỏ hơn 0.5, nên loại từng biến quan sát một

Rotated Component Matrix^a

| | Component | | |
|---|-----------|------|------|
| | 1 | 2 | 3 |
| I can be flexible with my time | .805 | | |
| I am satisfied with my salary | .704 | | |
| I am provided with computer | .664 | | |
| I do not have to wear uniform | .583 | | |
| My office is good and comfortable | .574 | | |
| My office have refresh training often | | .740 | |
| I receive adequate training to meet my current responsibilities. | | .739 | |
| I get good guidance and instruction from my boss | | .723 | |
| My supervisor response well to me when I have issues or concerns | .309 | .590 | |
| My boss like my work | | | .825 |
| I am contributing to improving the conditions of the communities I am working in. | | | .746 |
| I find my job to be motivating and I like to do it. | | | .698 |
| People in the office value my jobs | | | .488 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Quyết định cuối cùng

- Như vậy các biến quan sát đưa vào FA được rút gọn thành 3 nhân tố. Bạn cũng biết mỗi nhân tố gồm các biến nào.
- Bạn sẽ cần phân tích xem các biến này có ý nghĩa gì, từ đó đặt tên cho nhân tố. Tên cần đại diện cho các biến quan sát của nhân tố

Nhân số

- Trong Hộp Factor Analysis, chọn nút score, sau đó Regression để lưu lại các nhân số của nhân tố một cách tự động. Nhân số này đã được chuẩn hóa.
- Sau đó dùng nhân số này như một biến trong các phân tích tiếp theo
 - Hoặc tính nhân số của nhân tố bằng phương pháp trung bình cộng

Các phân tích sau khi thực hiện phân tích nhân tố

- Tiến hành các phân tích thông thường khác: ANOVA, regression
- Có thể dùng nhân số chuẩn hoặc tính nhân số

Các vấn đề với Factor Analysis

- Nó mang tính nghệ thuật nhiều hơn khoa học
 - Có nhiều phương pháp extraction (PCA, FA, etc.)
 - Có nhiều phương pháp rotation (Orthogonal, Oblique)
 - Số các factors được chọn? T
- Chính vì vậy mà đây là phương pháp rất hay!