Факультет Біомедичної інженерії

Національного технічного університету України

«Київський політехнічний інститут»

Кафедра біомедичної кібернетики

Дисципліна

«Інтелектуальний аналіз даних»

***Лабораторна робота № 3***

**Тема:** Проведення факторного аналізу показників функціонального стану людини

Виконав:

студент групи БС-81

Сєров О. В.

Перевірив:

викладач

доц. каф. БМК

Павлов В.А.

дата \_\_\_\_\_\_\_\_\_\_

підпис \_\_\_\_\_\_\_\_\_

Київ – 2020

**Мета роботи:** провести повноцінний аналіз показників функціонального стану людини за допомогою факторного аналізу.

**Завдання до роботи**

***Частина 1***. Проведення компонентного аналізу та виділення кількості головних факторів методом Головних Компонент в задачі аналізу показників функціонального стану людини. Одержання та обґрунтування початкового факторного рішення.

***Завдання.***

1.Завантажити дані змінних функціонального стану людини (крім змінної *VB*) – додаток 3. *Запропоновані для дослідження змінні відібрано шаговим регресійним алгоритмом у модель біологічного віку людини -змінна VB (вірніше VB - різниця біологічного та реального віку). Тобто ми будемо досліджувати структуру групування вхідних аргументів моделі біологічного віку людини, та будемо намагатися визначити (назвати) основні фактори що формують дану характеристику людини*.

2.Відмітити (опції) :

- дослідження проводимо на основі кореляційної матриці (*діапазони змінних у коваріаційній матриці надзвичайно різнорідні: висока імовірність поганої обумовленості матриці даних*).

-початкове рішення

- відбір компонент на основі величини власних значень кореляційної матриці

- неповернене рішення

- графік власних чисел

- КМО та ...

- Детермінант

- Антиобраз

- Метод обертання - ні

Проаналізувати: графік власних чисел, відсоток поясненої дисперсії, та значення КМО.

Послідовно виключити деяку кількість параметрів з найгіршим значенням MSE. Обґрунтувати кінцевий результат для одержаного початкового факторного рішення.

***Частина 2***. Відпрацювання навичок застосування методів обертання факторного рішення та інтерпретації одержаних факторів. Факторний аналізу показників функціонального стану людини.

***Завдання***

Додатково до встановлених опцій в частині 1. лабораторної роботи встановити опції:

Зберегти як змінні – регресія

Вивести матрицю коефіцієнтів значень факторів

Замінити середнім

Не виводити значення з низькими значеннями

Для одержаного початкового рішення з прийнятним значенням КМО становити послідовно опції - метод обертання: Варімакс, Прямий облімін, Квартимакс, Еквімакс, Промкс.

Приділити особливу увагу ортогональному обертанню. Для одержаного кінцевого рішення навести суттєву інтерпретацію факторів та їх аналітичні вирази. З найдених формулювань факторів, показників якості факторного аналізу та методу обертання початкового факторного рішення навести висновки для зробленого факторного аналізу показників функціонального стану людини.

**Теоретичні відомості**

**IK - Індекс Кердо**

Відомо, що відношення величини частоти серцевих скорочень і артеріального тиску в нормі залишається постійним. На основі цієї закономірності, для визначення стану вегетативної нервової системи використовують індекс Кердо́, який розраховують за формулою:



ІК – індекс Кердо;

AD – показник систолічного артеріального тиску;

d – показник діастолічного артеріального тиску;

Р – частота серцевих скорочень.

В нормі цей показник близький до одиниці. Позитивні значення індексу вказують на перевагу симпатичної регуляції тонусу судин, від'ємні - вплив вегетативної нервової системи на тонус.

**H - ЖЕЛ** - Зміна дихальних обсягів в результаті дії фізичних навантажень

Проба Генчі - Переносимость гіпоксії

**W** – вага

**Kk** – норма Ккал в день

**El** – ємність легень

**PSH** – Проба Штанге — функціональна проба з затримкою дихання під час вдиху, для оцінки дихальної функції.

**S** – вид спорту

**d2** – місце народження

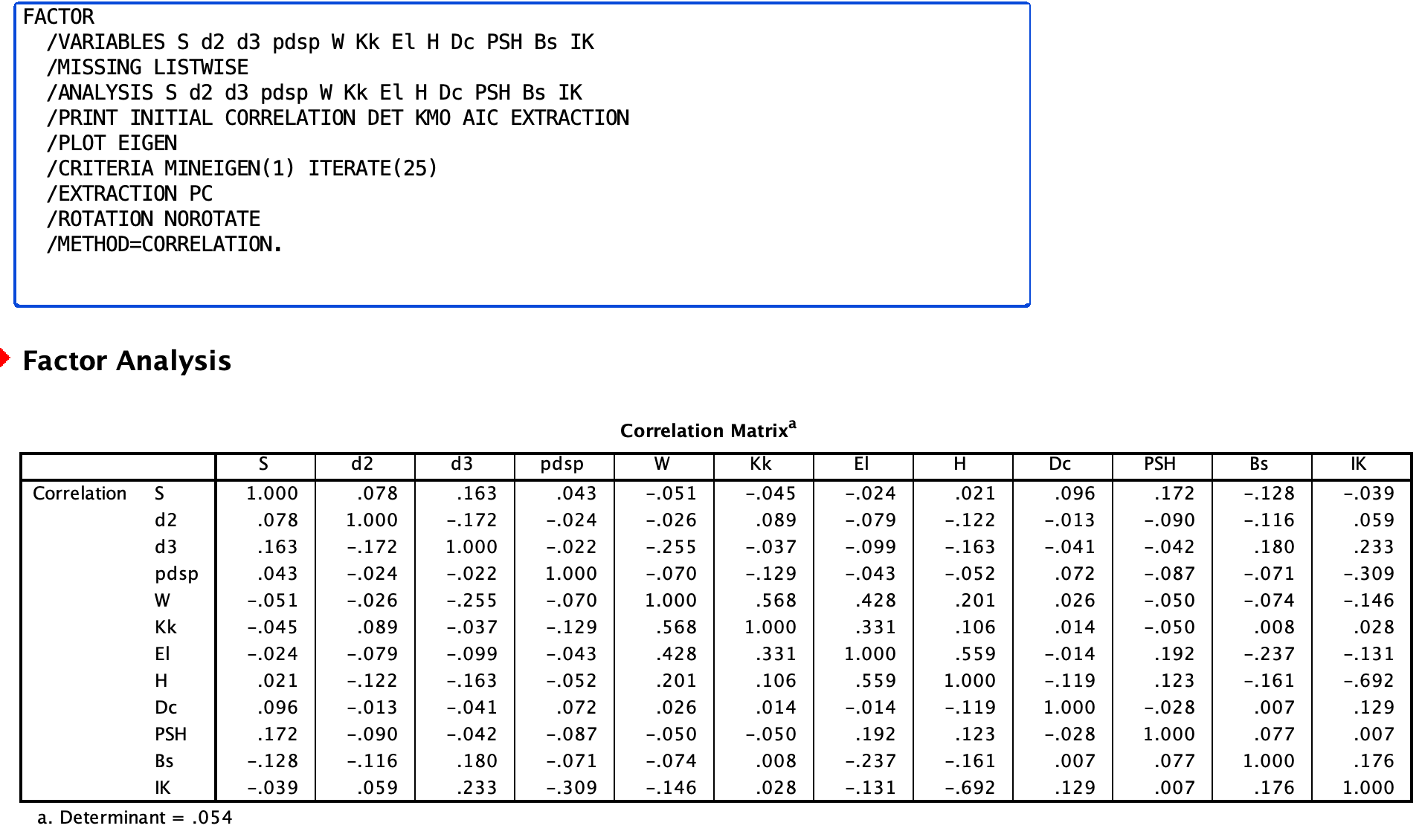
**Bs** – балансування

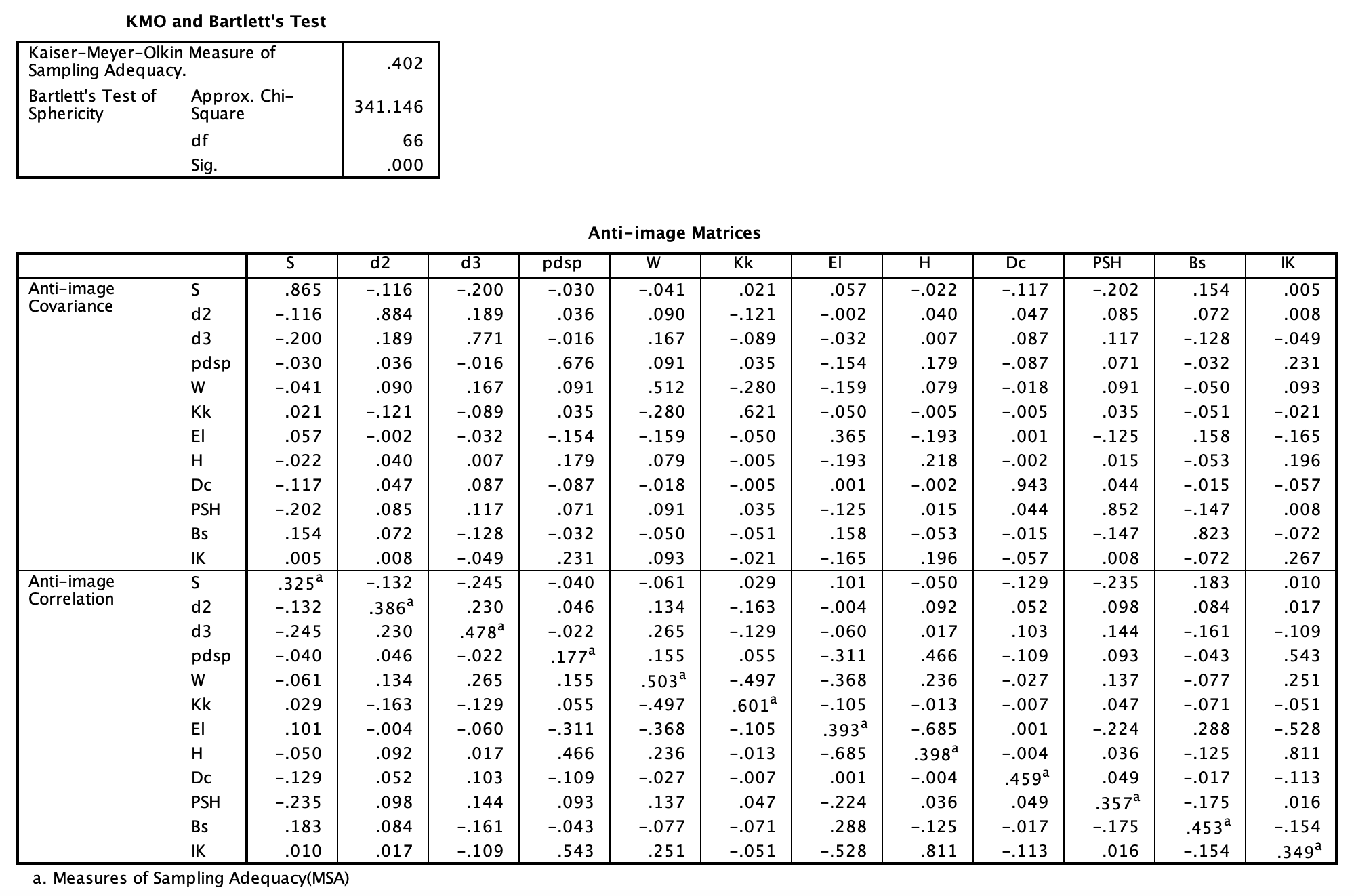
**Ds** – день менструального циклу у жінок

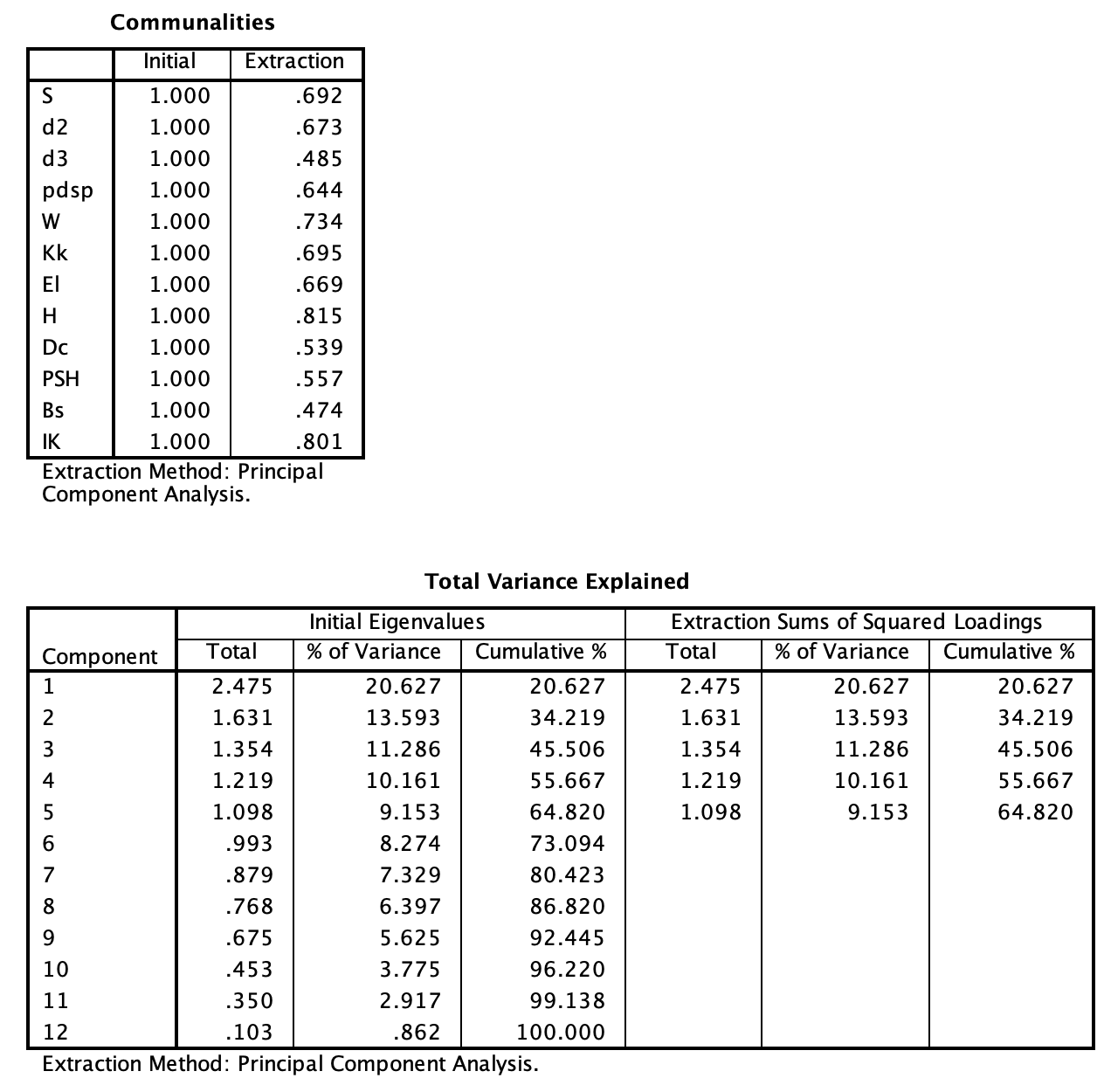
**pdsp** - АТДСП

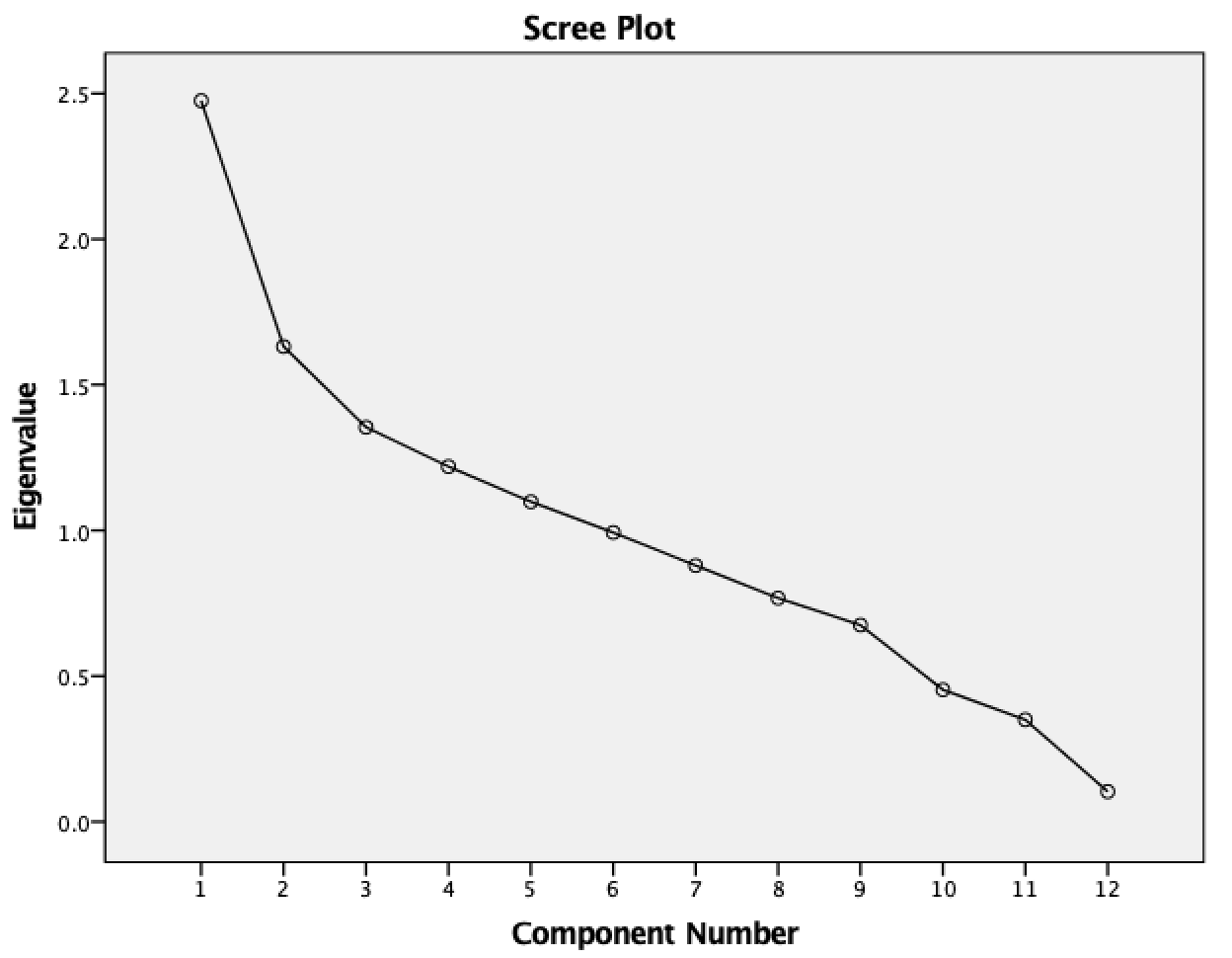
**d3** – рік народження

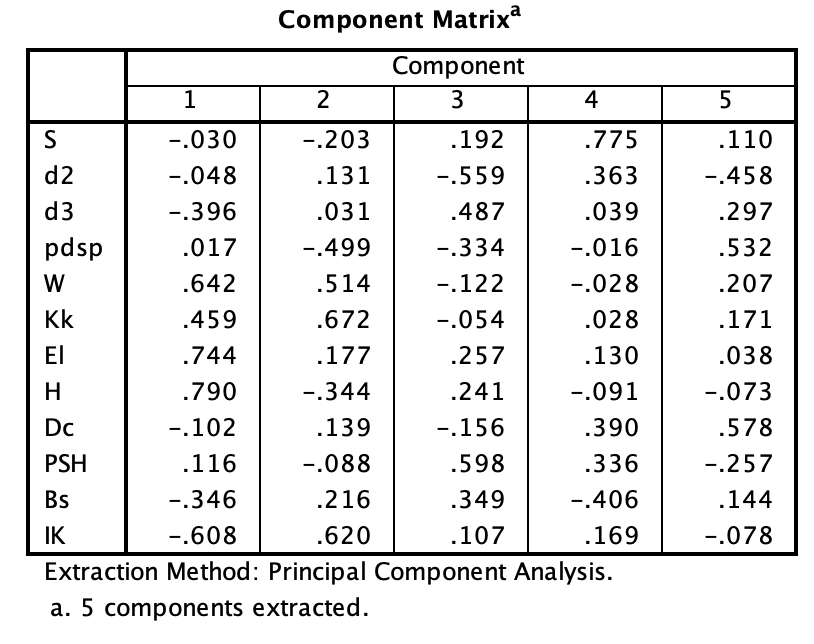
**Розрахунки та результати**







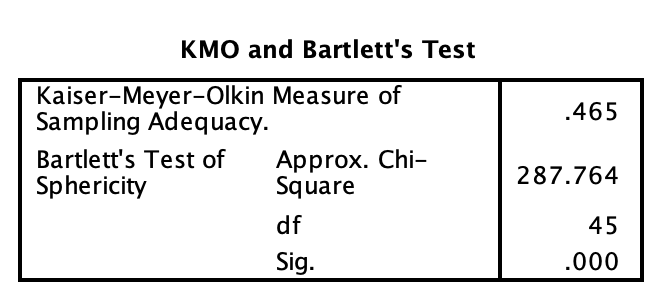




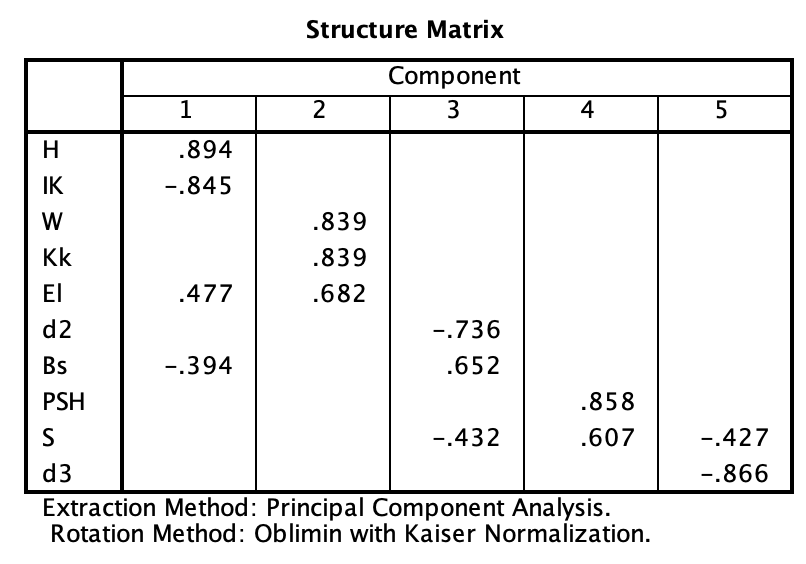
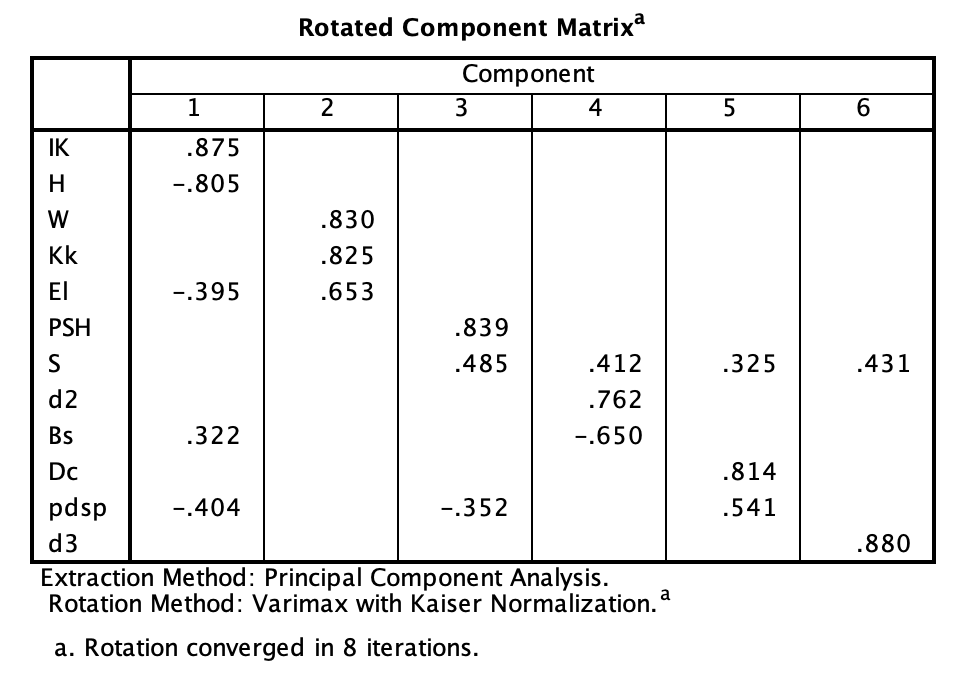
Міра адекватності вибірки Кайзера-Мейєра-Олкіна - це статистика, яка вказує на частку дисперсії у змінних, яка може бути спричинена основними факторами. Високі значення (близько 1,0), як правило, вказують на те, що факторний аналіз може бути корисним. Якщо значення менше 0,50, результати факторного аналізу, ймовірно, не будуть дуже корисними.

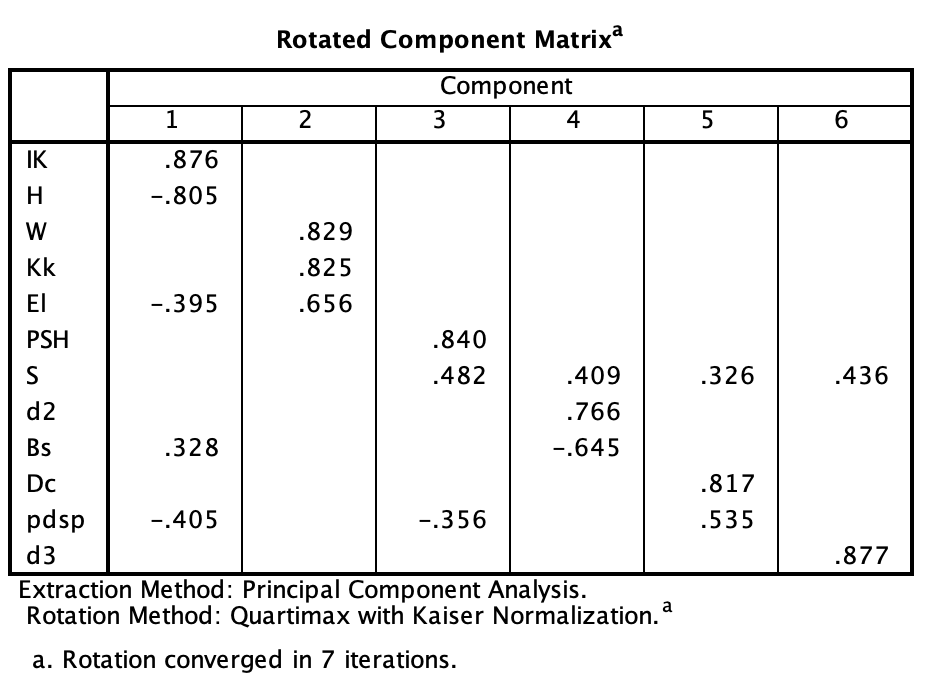
Наше значення – 0.402, що говорить про низьку якість отриманих факторів. Про це говорить також пологий графік.

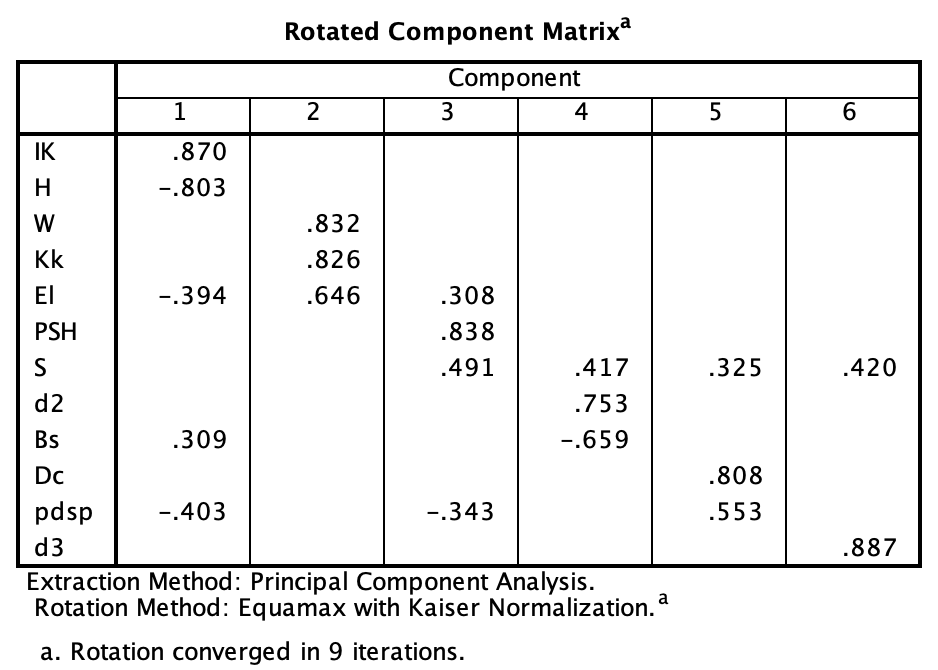
Після скорочення змінних до 10 ми отримали найоптимальніше значення КМО для цих даних:

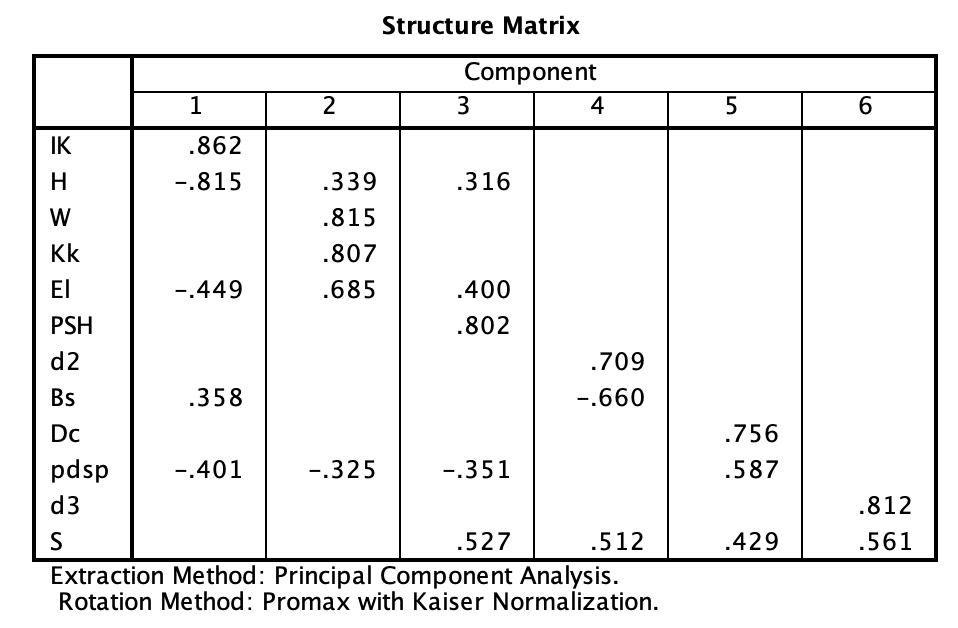


Додамо декілька опцій включаючи матрицю коефіцієнтів значень факторів та метод обертання почергово: Варімакс, Прямий облімін, Квартимакс, Еквімакс, Промакс.



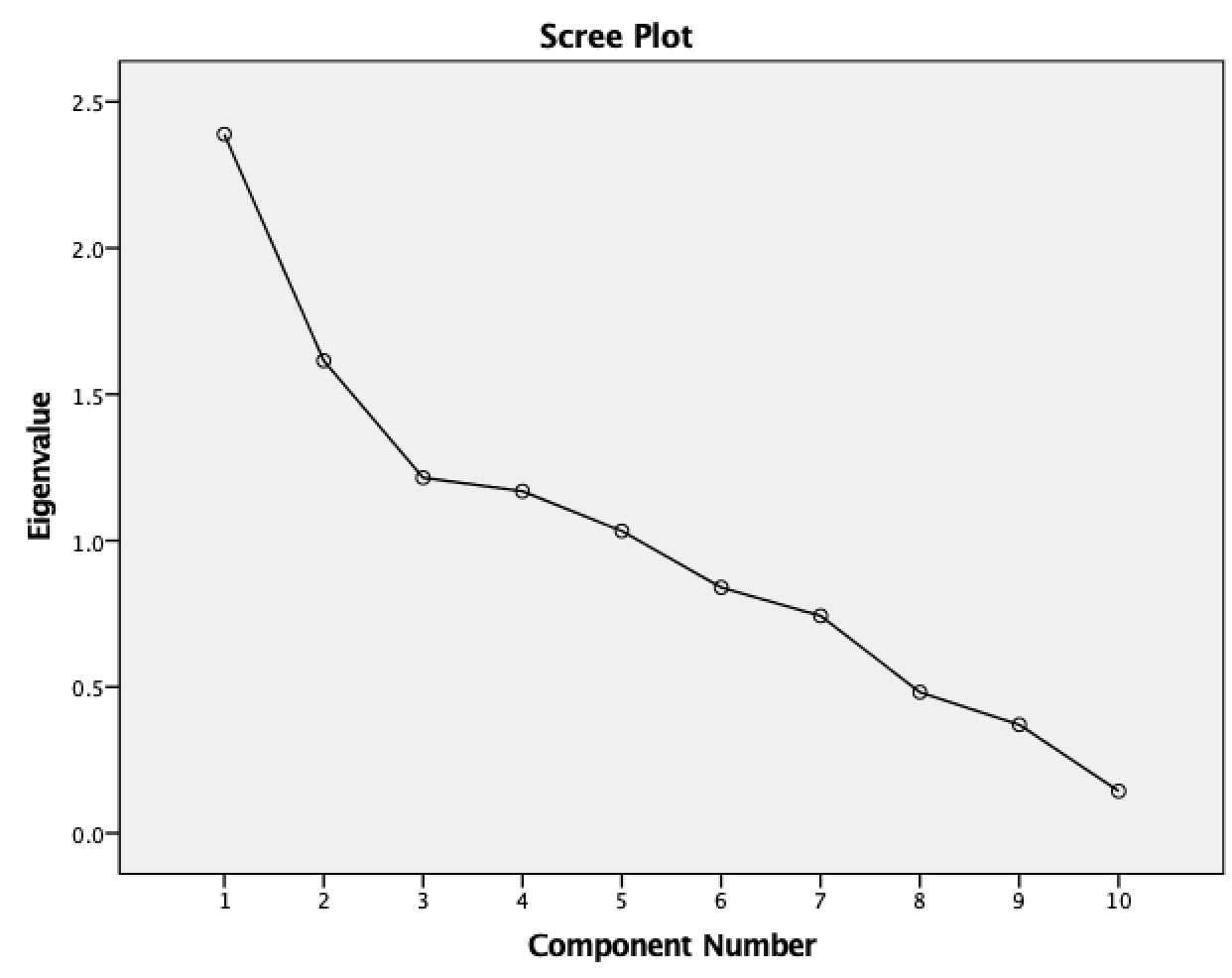






Скорегувавши дані та вибравши найнадійніший метод ортогонального обертання Варімакс проведемо інтерпретацію.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Rotated Component Matrixa** | | | | | |
|  | Component | | | | |
| 1 | 2 | 3 | 4 | 5 |
| H | -.882 |  |  |  |  |
| IK | .847 |  |  |  |  |
| Kk |  | .851 |  |  |  |
| W |  | .831 |  |  |  |
| El | -.435 | .651 |  |  |  |
| PSH |  |  | .867 |  |  |
| S |  |  | .592 | .449 | .374 |
| d2 |  |  |  | .720 | -.305 |
| Bs | .355 |  |  | -.664 |  |
| d3 |  |  |  |  | .877 |
| Extraction Method: Principal Component Analysis.  Rotation Method: Varimax with Kaiser Normalization. | | | | | |
| a. Rotation converged in 9 iterations. | | | | | |
|  | | | | | |
|  | | | | | |



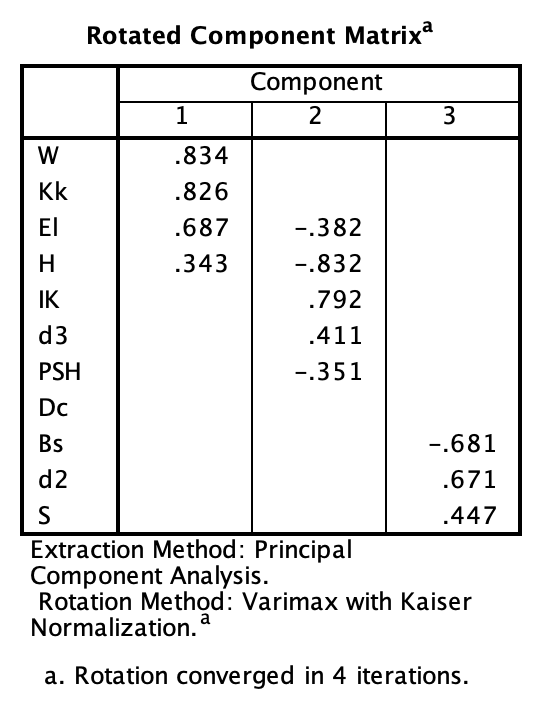
Аналізуючи отримані дані на розум спадає така інтерпретація:

1. Фізіологічні показники кисневого достатку організму (ЖЕЛ + ІК)
2. Запас енергії організму (Ккал за день, вага та ємність легень)
3. Спортивна підготовка/ефективність кисневого обміну
4. Цікаве об’єднання балансування та місця народження дає можливе об’єднання як вплив місцевості у дитинстві. Мається на увазі наявність активного способу життя в залежності від місцевості, наприклад у місті вестибулярний апарат дитини, на мою думку, менш розвинутий ніж у місцевості за водоймами та схилами для активностей.
5. Рік народження не скорочується, показуючи що від віку показники людини не міняються.

**Висновки**

Факторний аналіз показав прийнятні результати які можна інтерпретувати. Головною проблемою та недоліком цього методу полягає у суб’єктивності інтерпретування цих результатів, так як це можливо найважливіша частина аналізу. Без людини яка знається на даних з якими ми працюємо факторний аналіз може викликати труднощі чи привести до неправильних результатів.

З даними які у нас є по функціональному стану людини інтерпретації зводяться в цілому до 3 змінних:



Розмір організму, фізіологічних показників та цікавої змінної про ‘місце народження та балансування’. Цілком можливо що є більш точна інтерпретація по 5 факторам, але це потребує більш детального дослідження з залученням знавців медицини.