Principal Component Analysis and Exploratory Factor Analysis

Module 1: PCA

Karen Grace-Martin

Workshop Outline

Module 1: Principal Component Analysis Part 1

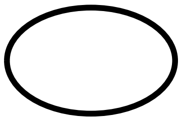
1. Principal Component vs. Exploratory Factor Analysis
2. Key Concepts and Aims
3. Initial Extraction of the First Component
4. Initial Extration of the Other Components
5. Determine the Number of Components to Retain
6. Example 2: Breastfeeding data

1. Principal Component vs. Exploratory Factor Analysis

Commonly Used Symbols of PCA, FA

Variables

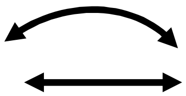
Observed Variable
- Item
- Indicator



Latent Variable

Relationships between Variables

Unidirectional Path



Correlation

Why do…

Principal Component Analysis?

Data Reduction:

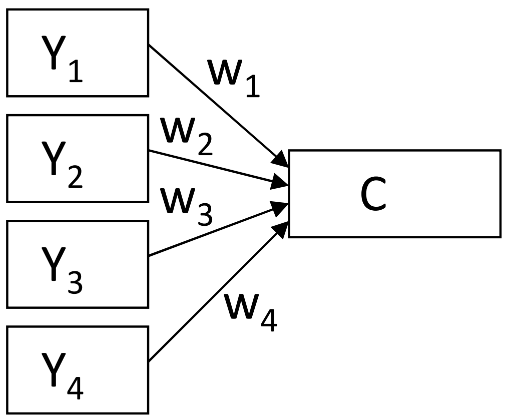
Replace a large number of variables with a smaller number that reflect most of the original data

Factor Analysis?

Test and Scale Construction:

Develop a measurement model to create “pure” measures of a construct

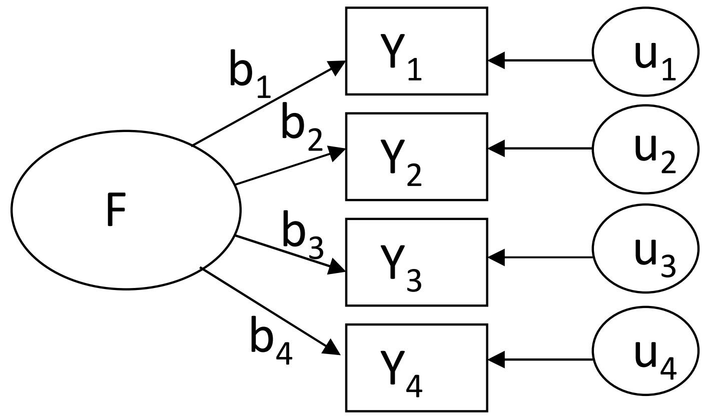
Principal Component Analysis



$$C = w_1(Y_1) + w_2(Y_2) + w_3(Y_3) + w_4(Y_4)$$

No error term

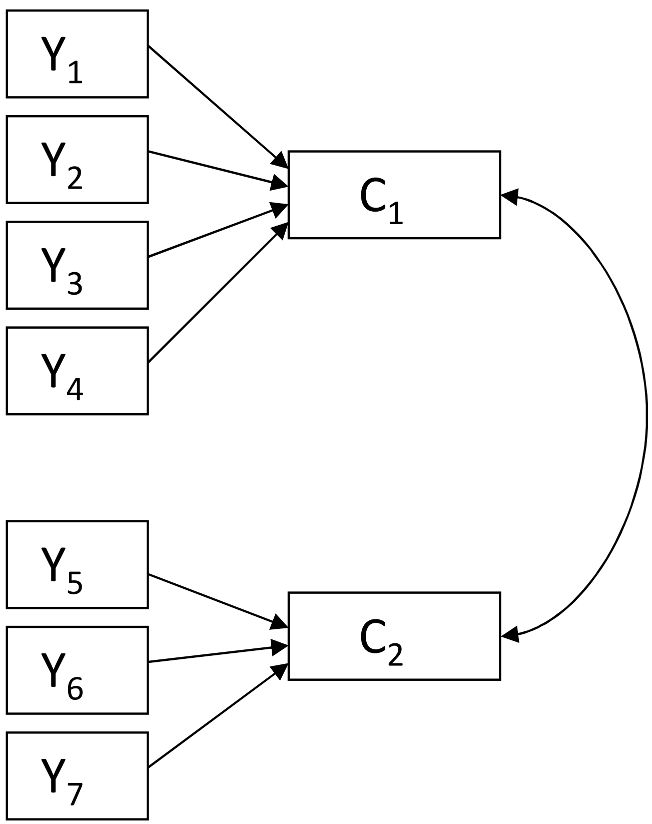
Factor Analysis



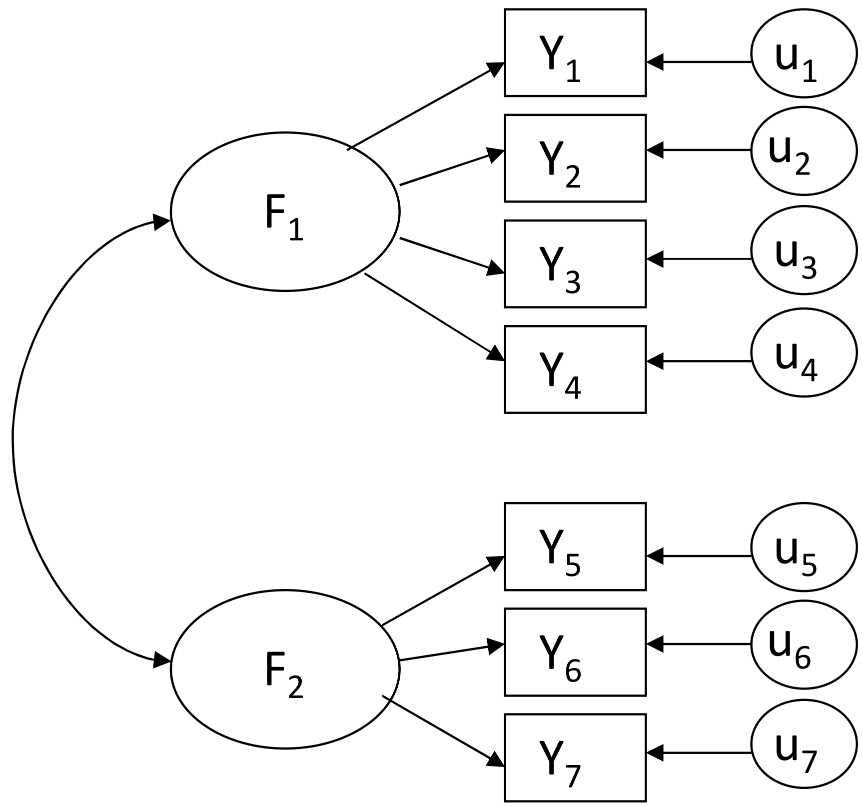
$$Y_1 = b_1{*}F + u_1$$
$$Y_2 = b_2{*}F + u_2$$
$$Y_3 = b_3{*}F + u_3$$
$$Y_4 = b_4{*}F + u_4$$

Each Y has a unique error term

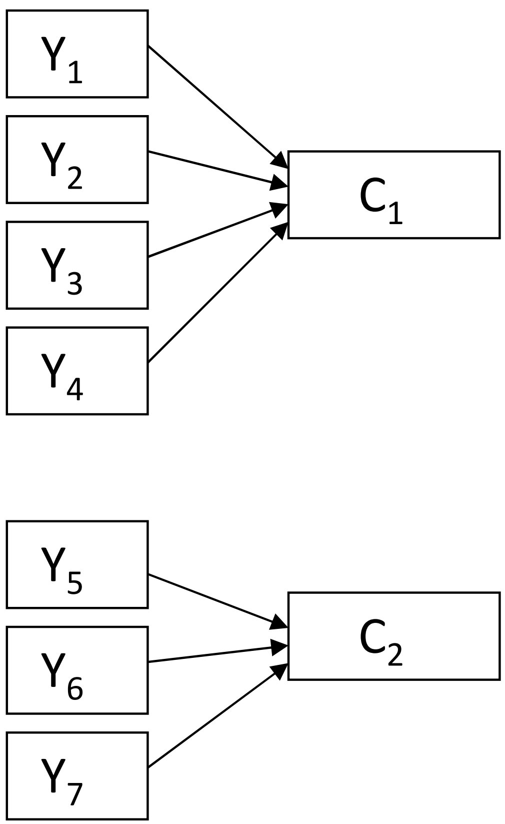
Principal Component Analysis



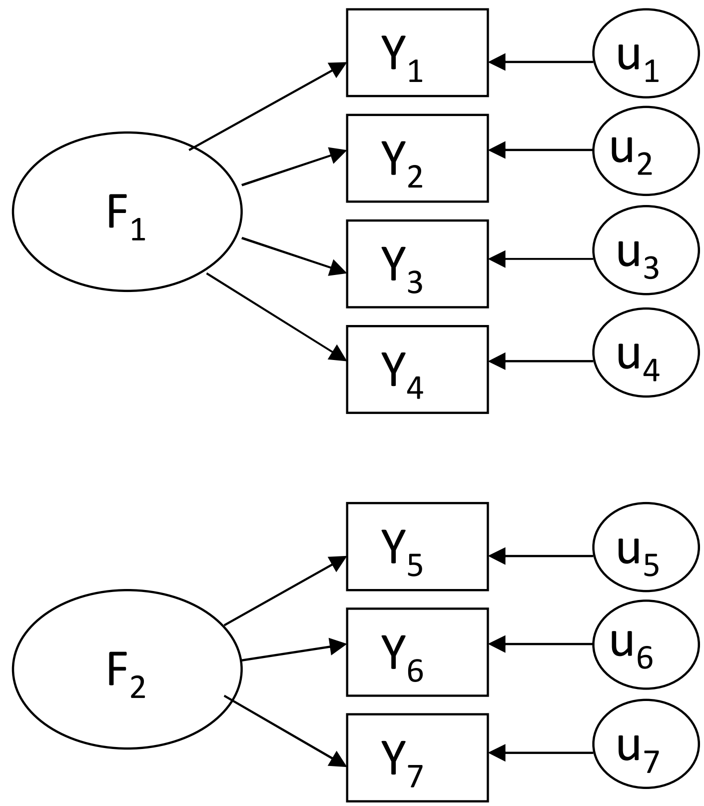
Factor Analysis

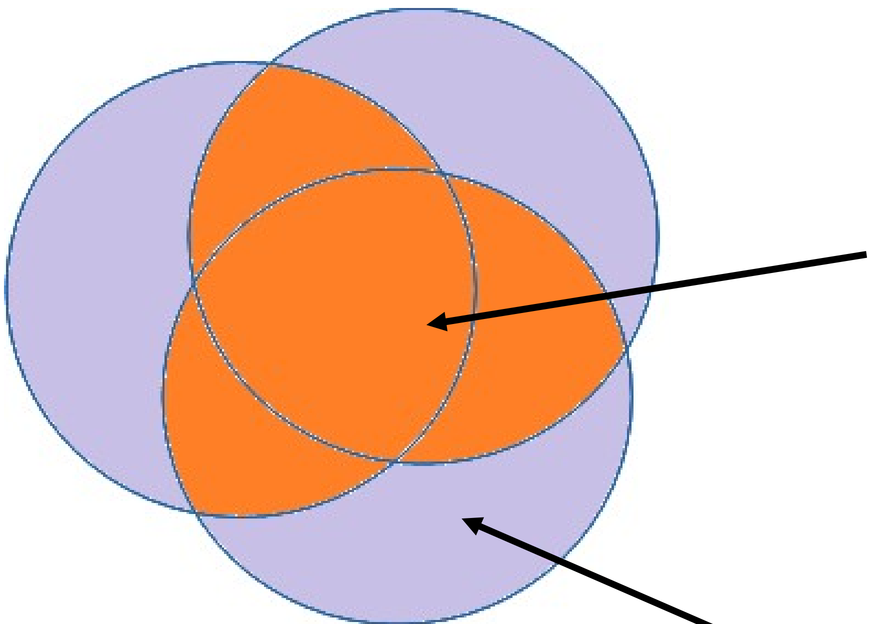


Principal Component Analysis



Factor Analysis





Common variance: h^2

Unique variance: $1 – h^2$

Steps are mostly the same

1. Initial Extraction of the Components/Factors
2. Determine the Number of Components/Factors to Retain
3. Rotation
4. Interpret the Rotated Solution
5. Create Component/Factor Scores

---

6. Summarizing the results in a Table
7. Preparing a formal description of the results for a paper

Principal Component Analysis or Factor Analysis?

Animal Data Set:
- Weight
- Total Sleep
- Predation
- Exposure During Sleep

Breastfeeding Data Set:
- Breastfeeding intention
- Self-Efficacy to Resist Formula

Caregivers:
- Mutual Communal Behavior Scale
- CES-D (Depression Scale)

Psych Data:
- Life Orientation (Optimism)
- Social Anxiety
- Attitude toward counseling

Principal Component Analysis – Example 1

Animal Data Set:

- Weight in kg (weight)
- Total Sleep in hours (totsleep)
- Predation on 1-5 index (predatn)
- Exposure During Sleep on 1-5 index (expos)

as those of clone C. The effectiveness of the clone B type of mechanism for protecting males against cannibalism, therefore, probably depends both on the frequency of encounters with cannibals which do attack them and on the voracity of such cannibals.

Males of clone C are readily attacked by cannibalistic female, especially campanulate, clonemates but have structural adaptations which protect them from being captured. A mechanism by which campanulates avoid attacking male clonemates may not have developed in clone C for several reasons. First, campanulates seem to be rare in this clone (*3*) and so may co-occur with male clonemates infrequently. Second, the great voracity of clone C campanulates might not be compatible with subtle feeding preferences.

It must be emphasized that male anticannibalism devices have been investigated only in these two quite distinct clones. The extent to which the development of the different protective mechanisms may be typical of, and effective within, the respective taxa—probably races or possibly closely related species (*7*)—is not known.

10. *HP-65 Stat Pac 1* (Hewlett-Packard, Cupertino, Calif., 1974).
11. The cruciform morphotype of clone B is unstable, rapidly transforming to the campanulate morphotype or to cruciform-campanulate intermediates (*3*), and was not studied.

12. Supported by NSF grant GB-31282, PHS grant 1 R01 HD 08210, and PHS research career development award K04-GM-70557. I thank M. Bean for expert technical assistance and T. M. Frost, J. R. Litton, Jr., P. L. Starkweather, and W. C. Kerfoot for improving the manuscript.

28 June 1976; revised 13 August 1976

Sleep in Mammals: Ecological and Constitutional Correlates

Abstract. *The interrelationships between sleep, ecological, and constitutional variables were assessed statistically for 39 mammalian species. Slow-wave sleep is negatively associated with a factor related to body size, which suggests that large amounts of this sleep phase are disadvantageous in large species. Paradoxical sleep is associated with a factor related to predatory danger, which suggests that large amounts of this sleep phase are disadvantageous in prey species.*

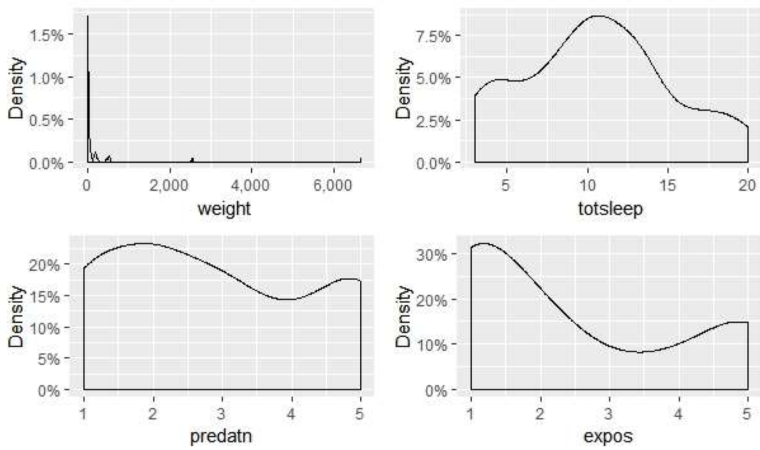
Every mammalian species studied in the laboratory spends at least some time asleep, and some species spend a major portion of their lives in this state (*1*). We may assume that sleep serves some (as yet unknown) biological function, but, if so, why do sleep requirements vary so much from species to species? By correlating sleep habits and other characteristics of species adapted to a wide variety of ecological niches it may be possible to clarify the significance of sleep in the life of mammals.

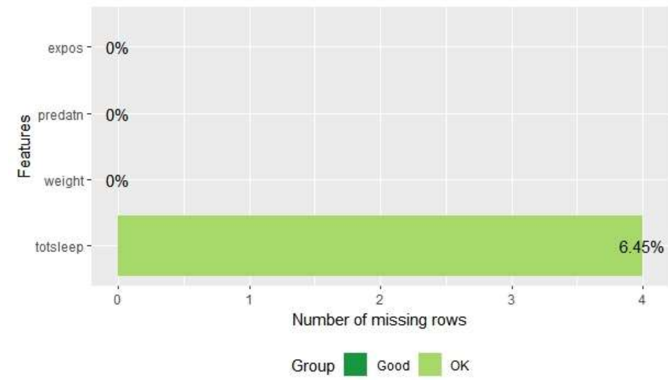
ily in the laboratory, and seem to need a great deal of sleep. Conversely, poor sleepers tend to sleep less and require long periods of adaptation to the laboratory before stable sleep habits are observed. These groups seem ecologically different in two ways. Predators (such as cats) are good sleepers, whereas species subject to heavy predation (such as rabbits) are poor sleepers. Second, species that sleep in reasonably secure places (such as bats) tend to sleep more than species that sleep in the open (such as

Source: Sleep in Mammals: Ecological and Constitutional Correlates" by Allison, T. and Cicchetti, D. (1976), Science, November 12, vol. 194, pp. 732-734.

Includes brain and body weight, life span, gestation time, time sleeping, and predation and danger indices for 62 mammals.

Principal Component Analysis – Example 1 - Descriptives





Variable	*M*	*SD*	*n*	SE_M	Skewness	Kurtosis
Weight	198.79	899.16	62	114.19	6.40	42.04
Totsleep	10.55	4.63	58	0.61	0.17	-0.65
Predatn	2.87	1.48	62	0.19	0.22	-1.32
Expos	2.42	1.60	62	0.20	0.66	-1.20

2. Key Concepts and Aims

Key Concepts

Review of Key General Concepts

- Correlation
- Sums of Squares
- Cross Products
- Covariance
- Variance
- Linear Combinations

Matrix Algebra Concepts in Bonus Video

- Matrix
- Vector
- Eigenvectors
- Eigenvalues

Introduction of Key Concepts in PCA and FA

- Extraction
- Loadings
- Communalities

The Aim of Principal Components

	weight	totsleep	predatn	expos
weight	1			
totsleep	-0.307	1		
predatn	0.059	-0.396	1	
expos	0.338	-0.642	0.618	1

Aim: rearrange the information in the correlation matrix by creating a set of components that are linear combinations of the original set of items.

Why: These components contain the variance in the original variables, but the variance is concentrated in the first few components.

Covariance, Variance, Correlation

Correlation Matrix **In generic units (std deviations)**

	weight	totsleep	predatn	expos
weight	1			
totsleep	-0.307	1		
predatn	0.059	-0.396	1	
expos	0.338	-0.642	0.618	1

$$Corr(X,Y) = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)\sqrt{Var(X)}\sqrt{Var(Y)}}$$

Variance/Covariance Matrix **In original units**

	weight	totsleep	predatn	expos
weight	808485.128			
totsleep	-1313.960	21.222		
predatn	78.981	-2.684	2.180	
expos	488.116	-4.544	1.465	2.575

$$Var(X) = \frac{\sum (X_i - \bar{X})(X_i - \bar{X})}{(n-1)}$$

$$Cov(X,Y) = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)}$$

What are Linear Combinations?

$Z = aX_1 + bX_2 + cX_3$

$W = (aX_1 + bX_2 + cX_3)/3$

Set of Weights

$$W = \frac{a}{3}X_1 + \frac{b}{3}X_2 + \frac{c}{3}X_3$$

Linear Combination

Variables

Rearranging Information Through Linear Combinations
Analogy: Egg Timer Puzzle

Goal: Boil an egg for 3 minutes

Tools: Two-minute timer
Seven-minute timer

Solution:
1. Start both timers: T=0
2. When two-minute timer ends, restart: T=2
3. When two-minute timer ends again, put egg in boiling water. Seven minute timer has 3 minutes left: T=4
4. When 7-minute timer ends, take egg out

S = run time for 7-minute timer

W = run time for 2-minute timer

C = cooking time

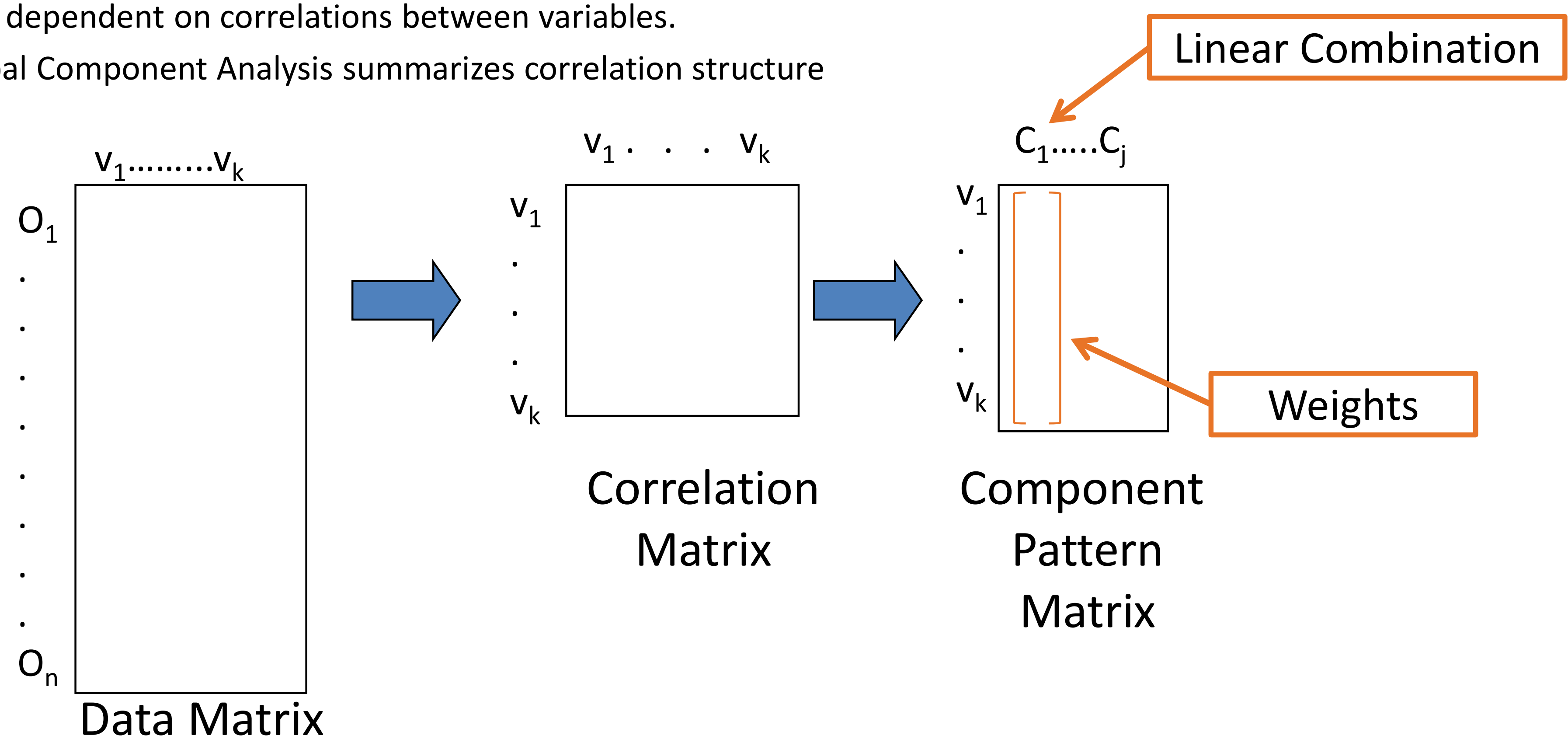
C = S – 2W

Linear Combination

Weights

The General Process

- Totally dependent on correlations between variables.
- Principal Component Analysis summarizes correlation structure



3. Initial Extraction of the First Component

Principal Component Analysis Steps

1. Initial Extraction of the Components
2. Determine the Number of Components to Retain
3. Rotation
4. Interpret the Rotated Solution
5. Create Component Scores

Initial Extraction of the Components

For each component, PCA computes two sets of values:

1.The Eigenvectors
(characteristic vectors) V_a

A column of weights that apply to each variable.

2.Corresponding Eigenvalues
(characteristic roots)

The total amount of variance explained by a component.

Component Loadings:

- Vector of correlations between the component and each variable
- Each element of the eigenvector * the square root of the corresponding eigenvalue.

Initial Extraction of the Components

Iterative algorithm designed to:

- Create Components: linear combinations of the original variables.
- Create same number of components as we have variables
- Find the one characteristic vector (eigenvector) that will give us the most variance explained (largest eigenvalue).

What you Get:

Total Variance Explained			
	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.247	56.180	56.180
2	.957	23.918	80.098
3	.524	13.096	93.195
4	.272	6.805	100.000

Eigenvalues

Component Loadings

Component Matrix[a]				
	Component			
	1	2	3	4
weight	.490	.816	.287	-.102
totsleep	-.816	-.052	.540	.198
predatn	.719	-.532	.386	-.228
expos	.908	-.066	.025	.413

Extraction Method: Principal Component Analysis.[a]

a. 4 components extracted.

What your software is doing to give you this:

Runs the following algorithm for each component, in order…..

Initial Extraction of the First Component (the Vector V_1)

Start with the correlation matrix:

$$
R = \begin{array}{l} \\ \textbf{weight} \\ \textbf{sleep} \\ \textbf{predation} \\ \textbf{expos} \end{array}
\begin{array}{c}
\begin{array}{rrrr} \textbf{weight} & \textbf{sleep} & \textbf{predation} & \textbf{expos} \end{array} \\
\begin{bmatrix} 1 & -0.307 & 0.052 & 0.356 \\ -0.307 & 1 & -0.396 & -0.642 \\ 0.052 & -0.396 & 1 & 0.603 \\ 0.356 & -0.642 & 0.603 & 1 \end{bmatrix}
\end{array}
$$

PC 1: Iteration 1

Initial Extraction of the Components

Step 1: Sum the coefficients in each row to create the first trial characteristic vector U_{a1}

$$
R = \begin{array}{l} \\ \text{weight} \\ \text{sleep} \\ \text{predation} \\ \text{expos} \end{array}
\begin{array}{c} \begin{array}{cccc} \text{weight} & \text{sleep} & \text{predation} & \text{expos} \end{array} \\
\begin{bmatrix} 1 & -0.307 & 0.052 & 0.356 \\ -0.307 & 1 & -0.396 & -0.642 \\ 0.052 & -0.396 & 1 & 0.603 \\ 0.356 & -0.642 & 0.603 & 1 \end{bmatrix} \end{array}
\quad U_{a1} = \begin{bmatrix} 1.101 \\ -0.345 \\ 1.259 \\ 1.317 \end{bmatrix}
$$

PC 1: Iteration 1

Initial Extraction of the Components

Step 2: Normalize U_{a1} to create V_{a1} (First Trial Characteristic Vector) by squaring and summing row totals, then dividing by the square root of the sums of squares

1 Square U's

R =

	weight	sleep	predation	expos
weight	1	-0.307	0.052	0.356
sleep	-0.307	1	-0.396	-0.642
predation	0.052	-0.396	1	0.603
expos	0.356	-0.642	0.603	1

U_{a1}=	Squares	Proportions of sqrt(SS) (V_{a1}=)
1.101	1.212	0.511
-0.345	0.119	-0.160
1.259	1.585	0.584
1.317	1.734	0.611

SS= 4.651

2 Sum

sqrt(SS)= 2.157

3 Square root

4 Divide U_{a1} by sqrt(SS)

PC 1: Iteration 1

Initial Extraction of the Components

Step 3: Produce the second trial vector, U_{a2} by multiplying R by V_{a1}

$$R = \begin{array}{c|cccc} & \text{weight} & \text{sleep} & \text{predation} & \text{expos} \\ \hline \text{weight} & 1 & -0.307 & 0.052 & 0.356 \\ \text{sleep} & -0.307 & 1 & -0.396 & -0.642 \\ \text{predation} & 0.052 & -0.396 & 1 & 0.603 \\ \text{expos} & 0.356 & -0.642 & 0.603 & 1 \end{array} * V_{a1} = \begin{bmatrix} 0.511 \\ -0.160 \\ 0.584 \\ 0.611 \end{bmatrix} = U_{a2} = \begin{bmatrix} 0.807 \\ -0.940 \\ 1.042 \\ 1.247 \end{bmatrix}$$

$$U_{a2} = \begin{bmatrix} 0.807 \\ -0.940 \\ 1.042 \\ 1.247 \end{bmatrix} \begin{array}{l} = 1*.511 -307*(-.16) + .052*.584 +.356*.611 \\ = -.307*.511 + 1*(-.16) - .396*.584 -.642*.611 \\ = .052*.511 -396*(-.16) + .1*.584 +.603*.611 \\ = .356*.511 -.642*(-.16) +.603*.584 + 1*.611 \end{array}$$

PC 1: Iteration 2

Initial Extraction of the Components

Step 4: Normalize the second trial vector U_{a2} to create V_{a2}

1 Square U's

		Squares		Proportions of sqrt(SS
U_{a2} =	0.807	0.652	V_{a2} =	0.395
	-0.940	0.884		-0.460
	1.042	1.086		0.510
	1.247	1.555		0.610

SS= 4.177

2 Sum

sqrt(SS)= 2.044

3 Square root

4 Divide U_{a2} by sqrt(SS)

PC 1: Iteration 2

Initial Extraction of the Components

Step 5: If V_{a2} differs from V_{a1}, create a third trial vector

Repeat these steps until subsequent characteristic vectors don't differ (convergence)

$$V_{a1} = \begin{bmatrix} 0.511 \\ -0.160 \\ 0.584 \\ 0.611 \end{bmatrix} \qquad V_{a2} = \begin{bmatrix} 0.395 \\ -0.460 \\ 0.510 \\ 0.610 \end{bmatrix}$$

PC 1: Iteration 3 -> i

Initial Extraction of the Components

Repeat these steps until subsequent characteristic vectors don't differ (convergence)

The final characteristic vector V_a is the eigenvector for principal component 1

The SS of the final characteristic vector is the eigenvalue for principal component 1

The component loadings are the elements of the eigenvector multiplied by the square root of the eigenvalue for PC 1.

- correlations between the component and the variable.

PC 1: Iteration 3 -> i

Eigenvalues, Eigenvectors, and Loadings

Eigenvalues

Total Variance Explained			
	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	$\boxed{2.247}$	56.180	56.180
2	.957	23.918	80.098
3	.524	13.096	93.195
4	.272	6.805	100.000

$$(.490)^2 + (-.816)^2 + (.719)^2 + (.908)^2 = \mathbf{2.247}$$

Loadings

Component Matrix	
	Component
	1
weight	.490
totsleep	-.816
predatn	.719
expos	.908

4. Initial Extraction of the Other Components

Initial Extraction of the Other Components

Now: Obtain the second principal component

Steps are the same, but instead of R, use the residual matrix

Initial Extraction of the Other Components

What the heck is the residual matrix?

Definition:
the correlation matrix with the first component partialed out

Computation:

- Create a new 4x4 matrix of Loadings Cross Products: the product of component loadings for each set of two variables
- Subtract those values from the original

Residual Matrix

Component Matrix[a]	
	Component
	1
weight	.490
totsleep	-.816
predatn	.719
expos	.908

Loadings Cross Products

$$\begin{array}{c} \begin{bmatrix} 0.490 & -0.816 & 0.719 & 0.908 \end{bmatrix} \end{array}$$

$$\begin{bmatrix} 0.490 \\ -0.816 \\ 0.719 \\ 0.908 \end{bmatrix} \quad \begin{array}{r|cccc} & \text{weight} & \text{sleep} & \text{predation} & \text{expos} \\ \hline \text{weight} & 0.240 & -0.400 & 0.352 & 0.445 \\ \text{sleep} & -0.400 & 0.666 & -0.587 & -0.741 \\ \text{predation} & 0.352 & -0.587 & 0.517 & 0.653 \\ \text{expos} & 0.445 & -0.741 & 0.653 & 0.824 \end{array}$$

Residual Matrix = R - CP

R
(original correlation matrix)

	weight	sleep	predation	expos
weight	1	-0.307	0.052	0.356
sleep	-0.307	1	-0.396	-0.642
predation	0.052	-0.396	1	0.603
expos	0.356	-0.642	0.603	1

Loadings Cross Products

	weight	sleep	predation	expos
weight	0.240	-0.400	0.352	0.445
sleep	-0.400	0.666	-0.587	-0.741
predation	0.352	-0.587	0.517	0.653
expos	0.445	-0.741	0.653	0.824

Residual Matrix

	weight	sleep	predation	expos
weight	0.760	0.093	-0.300	-0.089
sleep	0.093	0.334	0.191	0.099
predation	-0.300	0.191	0.483	-0.050
expos	-0.089	0.099	-0.050	0.176

Initial Extration of the Other Principal Components

Start with Residual Matrix

Run Interative Steps for each Component until convergence

Result is that component's eigenvector, eigenvalue, and loadings

Number of Components = Number of Variables

Each component extracts the maximum remaining variance from the variables

Eigenvalues, Eigenvectors, and Loadings

Total Variance Explained			
	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.247	56.180	56.180
2	.957	23.918	80.098
3	.524	13.096	93.195
4	.272	6.805	100.000

2.247
+ .957
+ .524
+ .272
= 4

Component Matrix[a]				
	Component			
	1	2	3	4
weight	.490	.816	.287	-.102
totsleep	-.816	-.052	.540	.198
predatn	.719	-.532	.386	-.228
expos	.908	-.066	.025	.413

Extraction Method: Principal Component Analysis.[a]

a. 4 components extracted.

Eigenvalues, Eigenvectors, and Loadings

Total Variance Explained

Component	Initial Eigenvalues Total	% of Variance	Cumulative %
1	2.247	56.180	56.180
2	.957	23.918	80.098
3	.524	13.096	93.195
4	.272	6.805	100.000

$$
\begin{aligned}
&(.490)^2 \\
&+ (-.816)^2 \\
&+ (.719)^2 \\
&+ (.908)^2 \\
&= \mathbf{2.247}
\end{aligned}
$$

Component Matrix[a]

	Component 1	2	3	4
weight	.490	.816	.287	-.102
totsleep	-.816	-.052	.540	.198
predatn	.719	-.532	.386	-.228
expos	.908	-.066	.025	.413

Extraction Method: Principal Component Analysis.[a]

a. 4 components extracted.

Eigenvalues, Eigenvectors, and Loadings

Total Variance Explained

	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.247	56.180	56.180
2	.957	23.918	80.098
3	.524	13.096	93.195
4	.272	6.805	100.000

$$(.908)^2 + (-.066)^2 + (.025)^2 + (.413)^2 = 1.000$$

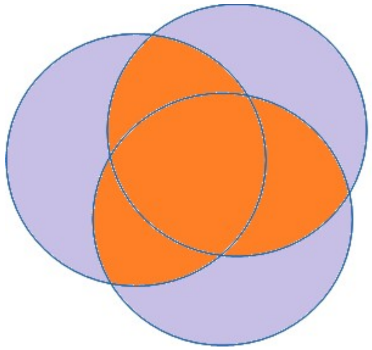
Component Matrix[a]

	Component			
	1	2	3	4
weight	.490	.816	.287	-.102
totsleep	-.816	-.052	.540	.198
predatn	.719	-.532	.386	-.228
expos	.908	-.066	.025	.413

Extraction Method: Principal Component Analysis.[a]

a. 4 components extracted.

Communalities



All Four Components

Communalities		
	Initial	Extraction
weight	1.000	1.000
totsleep	1.000	1.000
predatn	1.000	1.000
expos	1.000	1.000

Only One Component

Communalities		
	Initial	Extraction
weight	1.000	.240
totsleep	1.000	.666
predatn	1.000	.516
expos	1.000	.824

5. Determine the Number of Components to Retain

Principal Component Analysis Steps

1. Initial Extraction of the Components
2. Determine the Number of Components to Retain
3. Rotation
4. Interpret the Rotated Solution
5. Create Component Scores

Determine the Number of Components to Retain

Total Variance Explained

Component	Initial Eigenvalues Total	% of Variance	Cumulative %
1	2.247	56.180	56.180
2	.957	23.918	80.098
3	.524	13.096	93.195
4	.272	6.805	100.000

Component Matrix[a]

	Component 1	2	3	4
weight	.490	.816	.287	-.102
totsleep	-.816	-.052	.540	.198
predatn	.719	-.532	.386	-.228
expos	.908	-.066	.025	.413

Extraction Method: Principal Component Analysis.[a]

a. 4 components extracted.

Determine the Number of Components to Retain

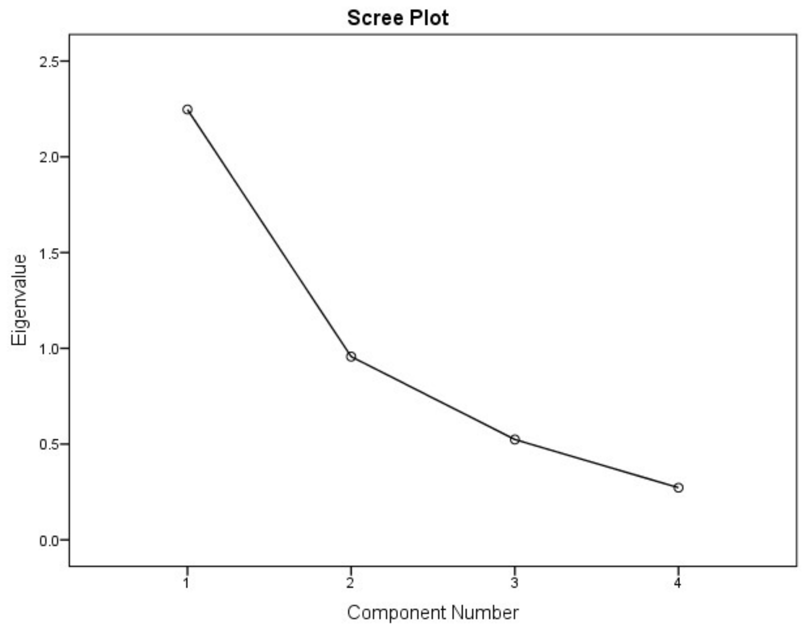
1. Kaiser criterion: Eigenvalue > 1
2. Percent variance accounted for
3. The Scree Test: Keep components before the break
4. **Parallel Analysis**
5. MAP Test
6. **The Interpretability criteria**

Determine the Number of Components to Retain

Total Variance Explained			
	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.247	56.180	56.180
2	.957	23.918	80.098
3	.524	13.096	93.195
4	.272	6.805	100.000

1. Kaiser criterion: Eigenvalue > 1
2. Percent variance accounted for

Determine the Number of Components to Retain



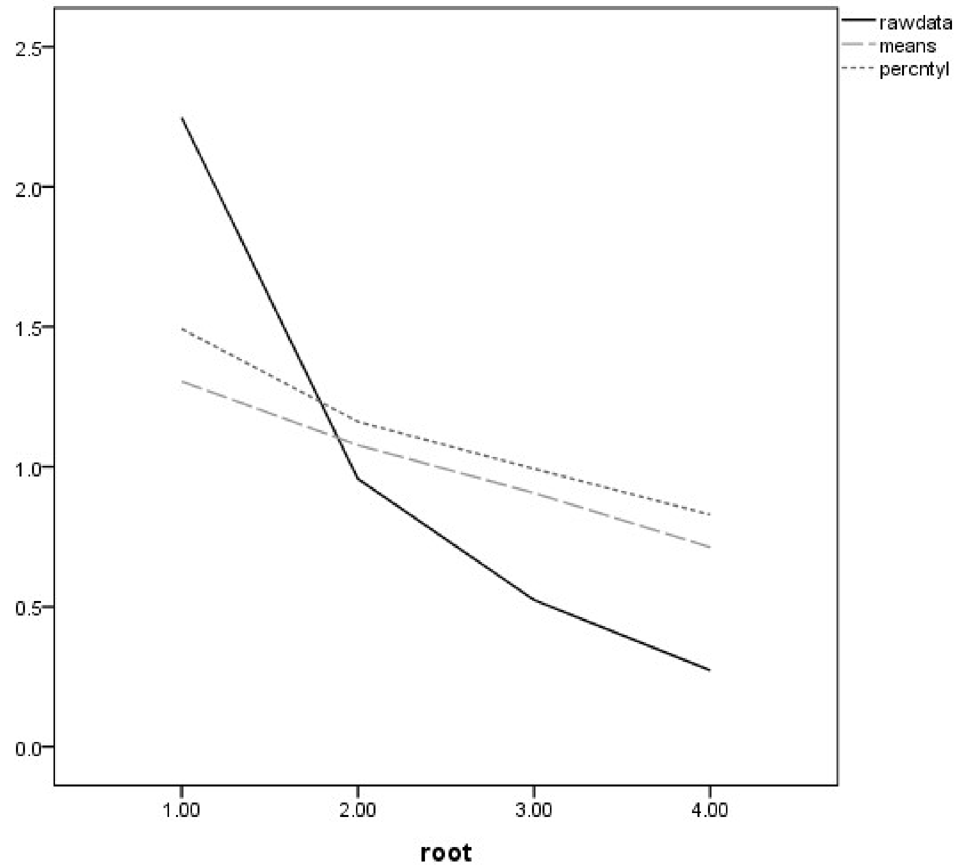
3. The Scree Test: Keep components *before* the break

Determine the Number of Components to Retain

4. Parallel Analysis

Steps:

1. Conduct a PCA on a random *nxp* data matrix
2. Repeat *k* times
3. Average the eigenvalues
4. Retain components with eigenvalues greater than the mean random eigenvalue



Determine the Number of Components to Retain

5. MAP Test

```
Velicer's Minimum Average Partial
 (MAP) Test:

Average Partial Correlations
                  squared     power4
          .0000    .1927      .0586
         1.0000    .1265      .0240
         2.0000    .4002      .2525
         3.0000   1.0000     1.0000
```

Steps:

1. Conduct a PCA
2. Partial the first PC out of the correlations
3. Compute the average squared coefficient in the off-diagonals
4. Repeat for each subsequent PC
5. Optimum number of components is the step with the lowest average squared partial correlation

```
The Number of Components According to the Original (1976) MAP Test is    1

The Number of Components According to the Revised (2000) MAP Test is    1
```

Determine the Number of Components to Retain

Use the Rotated Component Matrix

6. The Interpretability criteria

6.1 Minimum of three items with significant loadings on each retained component

6.2 The items that load on any component make sense

Component Matrix[a]

	Component			
	1	2	3	4
weight	.490	.816	.287	-.102
totsleep	-.816	-.052	.540	.198
predatn	.719	-.532	.386	-.228
expos	.908	-.066	.025	.413

Extraction Method: Principal Component Analysis.[a]

a. 4 components extracted.

6.3 Items loading on different components make sense

6.4 The rotated factor pattern has simple structure: Each variable has a relatively high loading (> .4) on only one component

Determine the Number of Components to Retain

Criterion	Number of Components Suggested
Kaiser	1
Total Variance Accounted for	2
Scree Plot	1
Parallel Analysis	1
MAP Test	1
Interpretability	1

Principal Component Analysis Steps

1. Initial Extraction of the Components
2. Determine the Number of Components to Retain
3. Rotation **Skip for today!**
4. Interpret the Rotated Solution
5. Create Component Scores **Skip for today!**

Interpret the Solution

Total Variance Explained			
	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.247	56.180	56.180
2	.957	23.918	80.098
3	.524	13.096	93.195
4	.272	6.805	100.000

Component Matrix[a]	
	Component
	1
weight	.490
totsleep	-.816
predatn	.719
expos	.908

A bit low

Options:

1. Split into two components.
2. Remove weight and see how that works.

Rerun without Weight and Compare

Total Variance Explained			
	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.247	56.180	56.180
2	.957	23.918	80.098
3	.524	13.096	93.195
4	.272	6.805	100.000

Total Variance Explained			
	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.101	70.026	70.026
2	.605	20.182	90.208
3	.294	9.792	100.000

Component Matrix	
	Component
	1
weight	.490
totsleep	-.816
predatn	.719
expos	.908

Component Matrix[a]	
	Component
	1
totsleep	-.812
predatn	.788
expos	.906

6. Example: Intent to Breastfeed Scale

The Variables

- HospBf Likely to breastfeed in the hospital
- HospF Likely to give formula in the hospital
- ONEmosBf Likely to breastfeed @ 1-mos
- ONEmosF Likely to give formula @ 1-mos
- FIVEmosBf Likely to breastfeed @ 5-mos
- FIVEmosF Likely to give formula @ 5-mos

Scaling
-2 extremely unlikely
-1 unlikely
0 neutral
1 likely
2 extremely likely

Steps for 2nd Example : Intent to Breastfeed Scale

1. Initial Extraction of the Components
2. Determine the Number of Components to Retain
3. Rotation **Skip for today!**
4. Interpret the Rotated Solution
5. Create Component Scores **Skip for today!**

Initial Extraction of the Components

Correlation Matrix

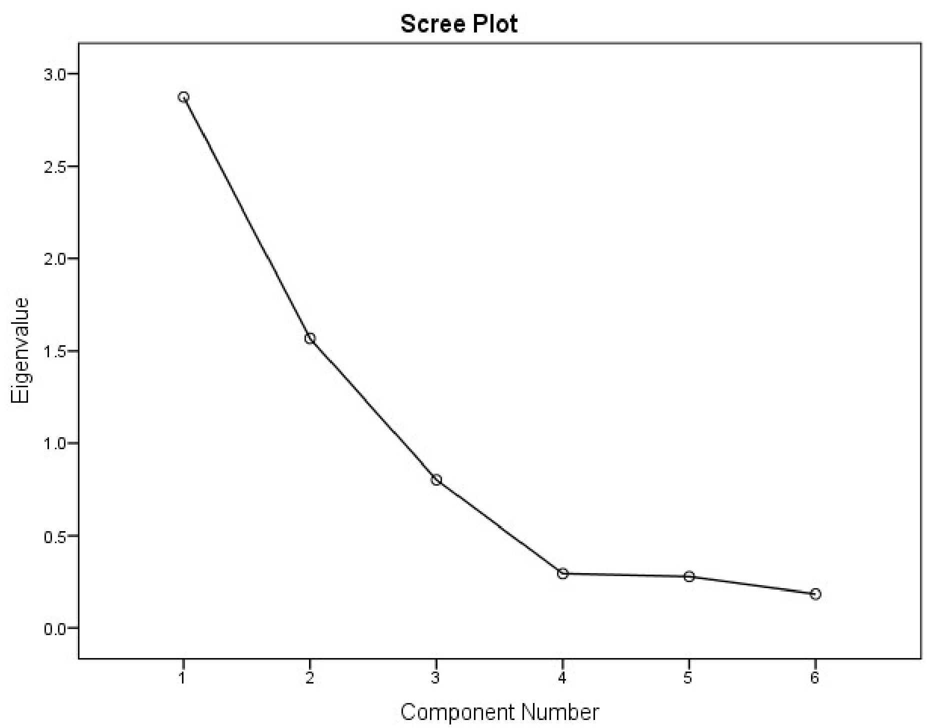
	HospBf	HospF	ONEmosBf	ONEmosF	FIVEmosBf	FIVEmosF
HospBf	1.000	.305	.713	.085	.370	.080
HospF	.305	1.000	.331	.752	.259	.553
ONEmosBf	.713	.331	1.000	.159	.490	.172
ONEmosF	.085	.752	.159	1.000	.164	.665
FIVEmosBf	.370	.259	.490	.164	1.000	.470
FIVEmosF	.080	.553	.172	.665	.470	1.000

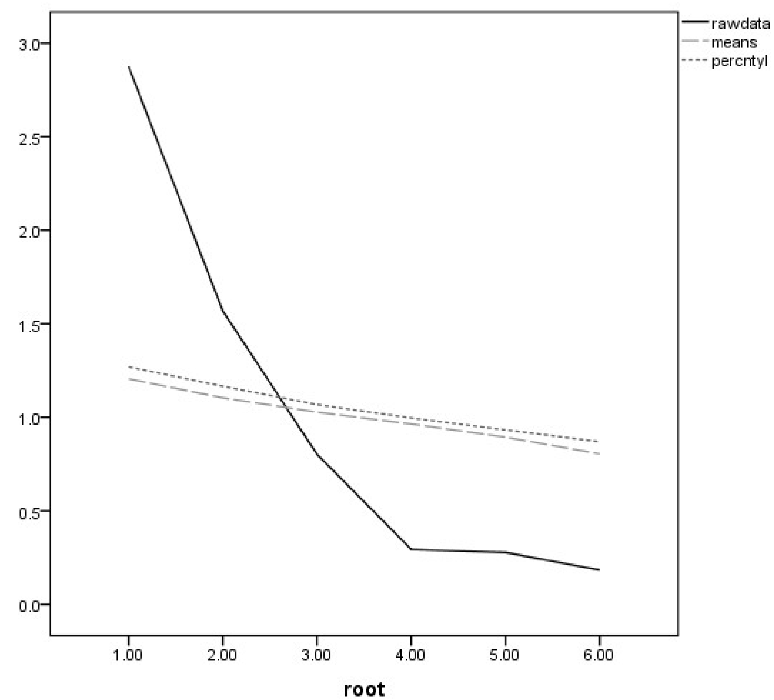
Initial Extraction of the Components

Total Variance Explained			
	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.876	47.926	47.926
2	1.568	26.128	74.054
3	.803	13.375	87.430
4	.293	4.889	92.319
5	.277	4.622	96.941
6	.184	3.059	100.000

Extraction Method: Principal Component Analysis.

Initial Extraction of the Components





Initial Extraction of the Components

Velicer's Minimum Average Partial (MAP) Test:

Eigenvalues

2.9148
1.5770
0.7787
0.2914
0.2582
0.1798

Average Partial Correlations

	squared	power4
0.0000	0.1925	0.0720
1.0000	0.2048	0.0640
2.0000	0.1636	0.0468
3.0000	0.2159	0.1140
4.0000	0.4133	0.3094
5.0000	1.0000	1.0000

The smallest average squared partial correlation is

0.1636

The smallest average 4rth power partial correlation is

0.0468

The Number of Components According to the Original (1976) MAP Test is

2.0000

The Number of Components According to the Revised (2000) MAP Test is

2.0000

Determine the Number of Components to Retain

The Interpretability criteria

1. Minimum of three items with high loadings on each retained component
2. The items that load on any component make sense
3. Items loading on different components make sense
4. The **rotated** factor pattern has simple structure: Each variable has a relatively high loading (> .4) on only one component

Rotated Component Matrix[a]

	Component	
	1	2
HospBf	.012	$\boxed{.882}$
HospF	$\boxed{.823}$	.265
ONEmosBf	.105	$\boxed{.903}$
ONEmosF	$\boxed{.920}$	.017
FIVEmosBf	.301	$\boxed{.645}$
FIVEmosF	$\boxed{.850}$	.128

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.[a]
a. Rotation converged in 3 iterations.

Determine the Number of Components to Retain

Criterion	Number of Components Suggested
Kaiser	2
Total Variance Accounted for	2
Scree Plot	3
Parallel Analysis	2
MAP Test	2
Interpretability	2

One Component Solution

Total Variance Explained			
	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	2.876	47.926	47.926
2	1.568	26.128	74.054
3	.803	13.375	87.430
4	.293	4.889	92.319
5	.277	4.622	96.941
6	.184	3.059	100.000

Extraction Method: Principal Component Analysis.

Component Matrix[a]	
	Component
	1
HospBf	.570
HospF	.803
ONEmosBf	.655
ONEmosF	.721
FIVEmosBf	.642
FIVEmosF	.738

Extraction Method: Principal Component Analysis.

In Module 2:

More about PCA:

1. Rotation for more than one component
2. Saving Component Scores