

“What is the annotated output for Principal Components Analysis in SAS?”

Authored by
stats writer

June 30, 2024

RECOMMENDED CITATION

stats writer (2024). “*What is the annotated output for Principal Components Analysis in SAS?*”. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=160708>

Principal Components Analysis (PCA) is a statistical technique used to reduce the dimensionality of a large dataset while retaining as much of the original information as possible. In SAS, the annotated output for PCA includes a detailed summary of the input data, the number of components used, and the percentage of variance explained by each component. The output also includes a scree plot, which visually displays the percentage of variance explained by each component. Additionally, SAS provides a table of factor loadings, which indicates the correlation between the original variables and the components. The annotated output for PCA in SAS is a comprehensive and informative summary of the results, making it a useful tool for data analysis and interpretation.

Principal Components Analysis | SAS Annotated Output

This page shows an example of a principal components analysis with footnotes explaining the output. The data used in this example were collected by Professor James Sidanius, who has generously shared them with us. You can download the data set here.

Overview: The "what" and "why" of principal components analysis

Principal components analysis is a method of data reduction. Suppose that you have a dozen variables that are correlated. You might use principal components analysis to reduce your 12 measures to a

few principal components.

In this example, you may be most interested in obtaining the component scores (which are variables that are added to your data set) and/or to look at

the dimensionality of the data. For example, if two components are extracted

and those two components accounted for 68% of the total variance, then we would

say that two dimensions in the component space account for 68% of the variance.

Unlike factor analysis, principal components analysis is not usually used to

identify underlying latent variables. Hence, the loadings onto the components

are not interpreted as factors in a factor analysis would be. Principal

components analysis, like factor analysis, can be preformed on raw data, as

shown in this example, or on a correlation or a covariance matrix. If raw data

is used, the procedure will create the original correlation matrix or covariance

matrix, as specified by the user. If the correlation matrix

is used, the variables are standardized and the total variance will equal the number of variables used in the analysis (because each standardized variable has a variance equal to 1). If the covariance matrix is used, the variables will remain in their original metric. However, one must take care to use variables whose variances and scales are similar. Unlike factor analysis, which analyzes the common variance, the original matrix in a principal components analysis analyzes the total variance. Also, principal components analysis assumes that each original measure is collected without measurement error.

In this example we have included many options, including the original correlation matrix and the scree plot. While you may not wish to use all of these options, we have included them here to aid in the explanation of the

analysis. We have also created a page of annotated output for a factor analysis that parallels this analysis. For general information regarding the similarities and differences between principal components analysis and factor analysis, please see our FAQ entitled What are some of the similarities and differences between principal components analysis and factor analysis?.

```
proc factor data = "d:m255_sas" corr scree ev method =
principal;
var item13 item14 item15 item16 item17 item18 item19
item20 item21 item22 item23 item24 ;
run;
```

Correlations

ITEM13 ITEM14 ITEM15

**ITEM13 INSTRUC WELL PREPARED 1.00000 0.66146
0.59999**

**ITEM14 INSTRUC SCHOLARLY GRASP 0.66146 1.00000
0.63460**

**ITEM15 INSTRUCTOR CONFIDENCE 0.59999 0.63460
1.00000**

**ITEM16 INSTRUCTOR FOCUS LECTURES 0.56626
0.50003 0.50535**

**ITEM17 INSTRUCTOR USES CLEAR RELEVANT
EXAMPLES 0.57687 0.55150 0.58664**

**ITEM18 INSTRUCTOR SENSITIVE TO STUDENTS
0.40898 0.43311 0.45707**

**ITEM19 INSTRUCTOR ALLOWS ME TO ASK
QUESTIONS 0.28632 0.32041 0.35869**

**ITEM20 INSTRUCTOR IS ACCESSIBLE TO STUDENTS
OUTSIDE CLASS 0.30418 0.31481 0.35568**

**ITEM21 INSTRUCTOR AWARE OF STUDENTS
UNDERSTANDING 0.47553 0.44896 0.50904**

**ITEM22 I AM SATISFIED WITH STUDENT
PERFORMANCE EVALUATION 0.33255 0.33313 0.36884**

**ITEM23 COMPARED TO OTHER INSTRUCTORS, THIS
INSTRUCTOR IS 0.56399 0.56461 0.58233**

**ITEM24 COMPARED TO OTHER COURSES THIS
COURSE WAS 0.45360 0.44281 0.43481**

Correlations

ITEM16 ITEM17 ITEM18

ITEM13 INSTRUC WELL PREPARED 0.56626 0.57687
0.40898

ITEM14 INSTRUC SCHOLARLY GRASP 0.50003 0.55150
0.43311

ITEM15 INSTRUCTOR CONFIDENCE 0.50535 0.58664
0.45707

ITEM16 INSTRUCTOR FOCUS LECTURES 1.00000
0.58649 0.40479

ITEM17 INSTRUCTOR USES CLEAR RELEVANT
EXAMPLES 0.58649 1.00000 0.55474

ITEM18 INSTRUCTOR SENSITIVE TO STUDENTS
0.40479 0.55474 1.00000

ITEM19 INSTRUCTOR ALLOWS ME TO ASK
QUESTIONS 0.33540 0.44930 0.62660

ITEM20 INSTRUCTOR IS ACCESSIBLE TO STUDENTS
OUTSIDE CLASS 0.31676 0.41682 0.52055

ITEM21 INSTRUCTOR AWARE OF STUDENTS
UNDERSTANDING 0.45245 0.59526 0.55417

ITEM22 I AM SATISFIED WITH STUDENT
PERFORMANCE EVALUATION 0.36255 0.44976 0.53609

ITEM23 COMPARED TO OTHER INSTRUCTORS, THIS
INSTRUCTOR IS 0.45880 0.61302 0.56950

ITEM24 COMPARED TO OTHER COURSES THIS
COURSE WAS 0.42967 0.52058 0.47382

Correlations

ITEM19 ITEM20 ITEM21

**ITEM13 INSTRUC WELL PREPARED 0.28632 0.30418
0.47553**

**ITEM14 INSTRUC SCHOLARLY GRASP 0.32041 0.31481
0.44896**

**ITEM15 INSTRUCTOR CONFIDENCE 0.35869 0.35568
0.50904**

**ITEM16 INSTRUCTOR FOCUS LECTURES 0.33540
0.31676 0.45245**

**ITEM17 INSTRUCTOR USES CLEAR RELEVANT
EXAMPLES 0.44930 0.41682 0.59526**

**ITEM18 INSTRUCTOR SENSITIVE TO STUDENTS
0.62660 0.52055 0.55417**

**ITEM19 INSTRUCTOR ALLOWS ME TO ASK
QUESTIONS 1.00000 0.44647 0.49921**

**ITEM20 INSTRUCTOR IS ACCESSIBLE TO STUDENTS
OUTSIDE CLASS 0.44647 1.00000 0.42479**

**ITEM21 INSTRUCTOR AWARE OF STUDENTS
UNDERSTANDING 0.49921 0.42479 1.00000**

**ITEM22 I AM SATISFIED WITH STUDENT
PERFORMANCE EVALUATION 0.48404 0.38297 0.50651**

ITEM23 COMPARED TO OTHER INSTRUCTORS, THIS INSTRUCTOR IS 0.44401 0.40962 0.59751

ITEM24 COMPARED TO OTHER COURSES THIS COURSE WAS 0.37383 0.35722 0.49977

Correlations

ITEM22 ITEM23 ITEM24

ITEM13 INSTRUC WELL PREPARED 0.33255 0.56399 0.45360

ITEM14 INSTRUC SCHOLARLY GRASP 0.33313 0.56461 0.44281

ITEM15 INSTRUCTOR CONFIDENCE 0.36884 0.58233 0.43481

ITEM16 INSTRUCTOR FOCUS LECTURES 0.36255 0.45880 0.42967

ITEM17 INSTRUCTOR USES CLEAR RELEVANT EXAMPLES 0.44976 0.61302 0.52058

ITEM18 INSTRUCTOR SENSITIVE TO STUDENTS 0.53609 0.56950 0.47382

ITEM19 INSTRUCTOR ALLOWS ME TO ASK QUESTIONS 0.48404 0.44401 0.37383

ITEM20 INSTRUCTOR IS ACCESSIBLE TO STUDENTS OUTSIDE CLASS 0.38297 0.40962 0.35722

ITEM21 INSTRUCTOR AWARE OF STUDENTS UNDERSTANDING 0.50651 0.59751 0.49977
ITEM22 I AM SATISFIED WITH STUDENT PERFORMANCE EVALUATION 1.00000 0.49317 0.44440
ITEM23 COMPARED TO OTHER INSTRUCTORS, THIS INSTRUCTOR IS 0.49317 1.00000 0.70464
ITEM24 COMPARED TO OTHER COURSES THIS COURSE WAS 0.44440 0.70464 1.00000

The table above was included in the output because we included the keyword `corr` on the `proc factor` statement. This table gives the correlations between the original variables (which are specified on the `var` statement). Before conducting a principal components analysis, you want to check the correlations between the variables. If any of the correlations are too high (say above .9), you may need to remove one of the variables from the analysis, as the two variables seem to be measuring the same thing. Another alternative would be to combine the variables in some

way (perhaps by taking the average). If the correlations are too low, say below .1, then one or more of the variables might load only onto one principal component (in other words, make its own principal component). This is not helpful, as the whole point of the analysis is to reduce the number of items (variables).

Initial Factor Method: Principal Components

Prior Communality Estimates: ONE

Eigenvalues of the Correlation Matrix: Total = 12
Average = 1

Eigenvaluea Differenceb Proportionc Cumulatived

1	6.24914661	5.01966832	0.5208	0.5208
2	1.22947829	0.51048923	0.1025	0.6232
3	0.71898906	0.10585957	0.0599	0.6831
4	0.61312949	0.05196458	0.0511	0.7342
5	0.56116491	0.05817383	0.0468	0.7810
6	0.50299107	0.03172750	0.0419	0.8229
7	0.47126357	0.08244834	0.0393	0.8622

8 0.38881523 0.02091149 0.0324 0.8946
9 0.36790373 0.03970330 0.0307 0.9252
10 0.32820043 0.01082277 0.0274 0.9526
11 0.31737767 0.06583773 0.0264 0.9790
12 0.25153994 0.0210 1.0000

2 factors will be retained by the MINEIGEN criterion.

a. Eigenvalue - This column contains the eigenvalues. The first component will always account for the most variance (and hence have the highest eigenvalue), and the next component will account for as much of the left over variance as it can, and so on. Hence, each successive component will account for less and less variance.

b. Difference - This column gives the differences between the current and the next eigenvalue. For example, $6.24 - 1.22 = 5.02$.

This gives you a sense of how much change there is in the eigenvalues from one

component to the next.

c. Proportion - This column gives the proportion of variance

accounted for by each component. In this example, the first component

accounts for just over half of the variance (approximately 52%).

d. Cumulative - This column sums up to proportion column, so

that you can see how much variance is accounted for by, say, the first five components, .7810.

Initial Factor Method: Principal Components

Scree Plot of Eigenvalues

|
7 +
|
|
|
|
|
| 1

6 +

|

|

|

|

|

5 +

|

|

E |

i |

g |

e 4 +

n |

v |

a |

l |

u |

e 3 +

s |

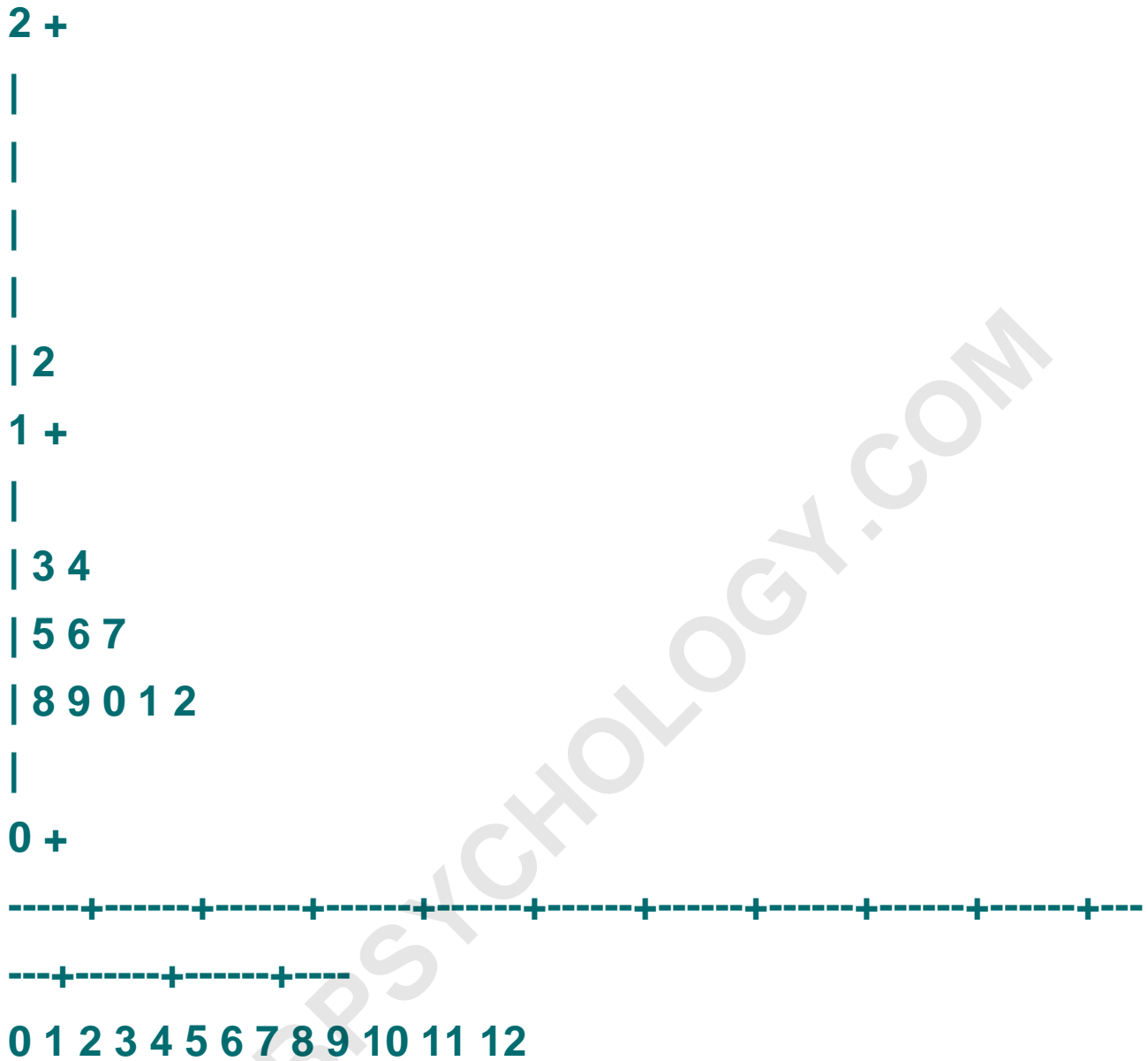
|

|

|

|

ARABPSYCHOLOGY.COM



Number

Initial Factor Method: Principal Components

The scree plot graphs the eigenvalue against the component number. You can see these values in the first two columns of the table

immediately above. From the third component on, you can see that the line is almost flat, meaning the each successive component is accounting for smaller and smaller amounts of the total variance. In general, we are interested in keeping only those principal components whose eigenvalues are greater than 1. Components with an eigenvalue of less than 1 account for less variance than did the original variable (which had a variance of 1), and so are of little use. Hence, you can see that the point of principal components analysis is to redistribute the variance in the correlation matrix (using the method of eigenvalue decomposition) to redistribute the variance to first components extracted.

Eigenvectors

1e 2e

ITEM13 INSTRUC WELL PREPARED 0.29093 -0.40510

ITEM14 INSTRUC SCHOLARLY GRASP 0.28953 -0.36765

ITEM15 INSTRUCTOR CONFIDENCE 0.29851 -0.27789
ITEM16 INSTRUCTOR FOCUS LECTURES 0.27406
-0.25376
ITEM17 INSTRUCTOR USES CLEAR RELEVANT
EXAMPLES 0.32261 -0.09492
ITEM18 INSTRUCTOR SENSITIVE TO STUDENTS
0.30207 0.33002
ITEM19 INSTRUCTOR ALLOWS ME TO ASK
QUESTIONS 0.25641 0.44823
ITEM20 INSTRUCTOR IS ACCESSIBLE TO STUDENTS
OUTSIDE CLASS 0.23709 0.34083
ITEM21 INSTRUCTOR AWARE OF STUDENTS
UNDERSTANDING 0.30536 0.12133
ITEM22 I AM SATISFIED WITH STUDENT
PERFORMANCE EVALUATION 0.26057 0.32871
ITEM23 COMPARED TO OTHER INSTRUCTORS, THIS
INSTRUCTOR IS 0.32768 -0.03634
ITEM24 COMPARED TO OTHER COURSES THIS
COURSE WAS 0.28550 0.00421

e. Eigenvectors - These columns give the eigenvectors for each variable in the principal components analysis. An eigenvector is a linear

combination of the original variables. The two components that have been extracted are orthogonal to one another, and they can be thought of as weights.

These weights are multiplied by each value in the original variable, and those values are then summed up to yield the eigenvector. The eigenvectors tell you about the strength of relationship between the variables and the components.

Factor Pattern

Factor1 Factor2

ITEM13 INSTRUC WELL PREPARED	0.72729	-0.44919
ITEM14 INSTRUC SCHOLARLY GRASP	0.72378	-0.40766
ITEM15 INSTRUCTOR CONFIDENCE	0.74622	-0.30813
ITEM16 INSTRUCTOR FOCUS LECTURES	0.68511	-0.28137
ITEM17 INSTRUCTOR USES CLEAR RELEVANT EXAMPLES	0.80647	-0.10525
ITEM18 INSTRUCTOR SENSITIVE TO STUDENTS	0.75512	0.36593
ITEM19 INSTRUCTOR ALLOWS ME TO ASK		

QUESTIONS 0.64098 0.49700
ITEM20 INSTRUCTOR IS ACCESSIBLE TO STUDENTS
OUTSIDE CLASS 0.59269 0.37792
ITEM21 INSTRUCTOR AWARE OF STUDENTS
UNDERSTANDING 0.76335 0.13454
ITEM22 I AM SATISFIED WITH STUDENT
PERFORMANCE EVALUATION 0.65138 0.36448
ITEM23 COMPARED TO OTHER INSTRUCTORS, THIS
INSTRUCTOR IS 0.81914 -0.04029
ITEM24 COMPARED TO OTHER COURSES THIS
COURSE WAS 0.71371 0.00467

f. Factor1 and Factor2 - This is the component matrix. This table contains component loadings, which are the correlations between the variable and the component. Because these are correlations, possible values range from -1 to +1. The columns under these headings are the principal components that have been extracted. As you can see, two components were extracted (the two components that had an eigenvalue greater than 1). You usually do not try to interpret the components the way

that you would factors that have been extracted from a factor analysis. Rather, most people are interested in the component scores, which are used for data reduction (as opposed to factor analysis where you are looking for underlying latent continua).

Variance Explained by Each Factor

Factor1 Factor2

6.2491466 1.2294783

Final Commuality Estimates: Total = 7.478625

ITEM13 ITEM14 ITEM15 ITEM16 ITEM17 ITEM18

**0.73071411 0.69004215 0.65179276 0.54854615
0.66147090 0.70412023**

Initial Factor Method: Principal Components

ITEM19 ITEM20 ITEM21 ITEM22 ITEM23 ITEM24

0.65786784 0.49410612 0.60081090 0.55713785

0.67261205 0.50940384

ARABPSYCHOLOGY.COM