Using our Data that was available from 1991 onwards we ran our first principal component analysis (PCA). The monthly data was used to increase the amount of observations to get more precise results.

[1] "Group.1" "AluIndex" "AluPrice" "BBIndex" "BBEnergyIndex" "BBIndustrialIndex" "CopperPrice"

[8] "CopperIndex" "GasPrice" "LeadIndex" "LeadPrice" "NickelIndex" "NickelPrice" "OilPrice"

[15] "Oil2Price" "GoldPrice" "PlatinumPrice" "Prec.met.Index" "SPIndex" "SilverPrice" "SP.AgriIndex"

[22] "SP.Agri.LiveIndex" "SP.AluminiumIndex" "SP.CopperIndex" "SP.crude.oilIndex" "SP.EnergyIndex" "SP.GoldIndex" "SP.LivestockIndex"

[29] "SP.prec.metIndex" "SPIndustrialIndex" "TinIndex" "TinPrice" "ZincIndex" "ZincPrice"

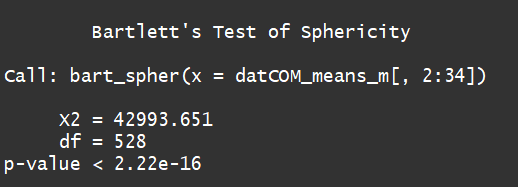
Firstly, we must run Bartlett's Test of Sphericity to find out if the data can be used for PCA. If this test reveals significant results the correlation between the variables in the data is high enough for PCA.

Call: bart\_spher(x = datCOM\_means\_m[, 2:34])

X2 = 42993.651

df = 528

p-value < 2.22e-16



From this test we can see that we can use PCA to reduce the dimensionality of the data.

For a final check we use the Kaiser-Meyer-Olkin Statistics to find out the KMOS-Criterion for PCA, for our data. This test shows us if the sample size is large enough. If the value is above 0.8 one can assume that the amount of observations is large enough. (when we used quarterly data this was not the case).

Call: KMOS(x = datCOM\_means\_m[, 2:34])

Measures of Sampling Adequacy (MSA):

AluIndex AluPrice BBIndex BBEnergyIndex BBIndustrialIndex CopperPrice CopperIndex GasPrice

0.9267828 0.9378160 0.8808483 0.6918151 0.9087768 0.9407134 0.7597650 0.8534875

LeadIndex LeadPrice NickelIndex NickelPrice OilPrice Oil2Price GoldPrice PlatinumPrice

0.7627126 0.7990782 0.8469011 0.7395728 0.9137082 0.9163716 0.8540521 0.9684933

Prec.met.Index SPIndex SilverPrice SP.AgriIndex SP.Agri.LiveIndex SP.AluminiumIndex SP.CopperIndex SP.crude.oilIndex

0.8729387 0.8777560 0.8743448 0.8367184 0.8293080 0.8538422 0.8945067 0.9155571

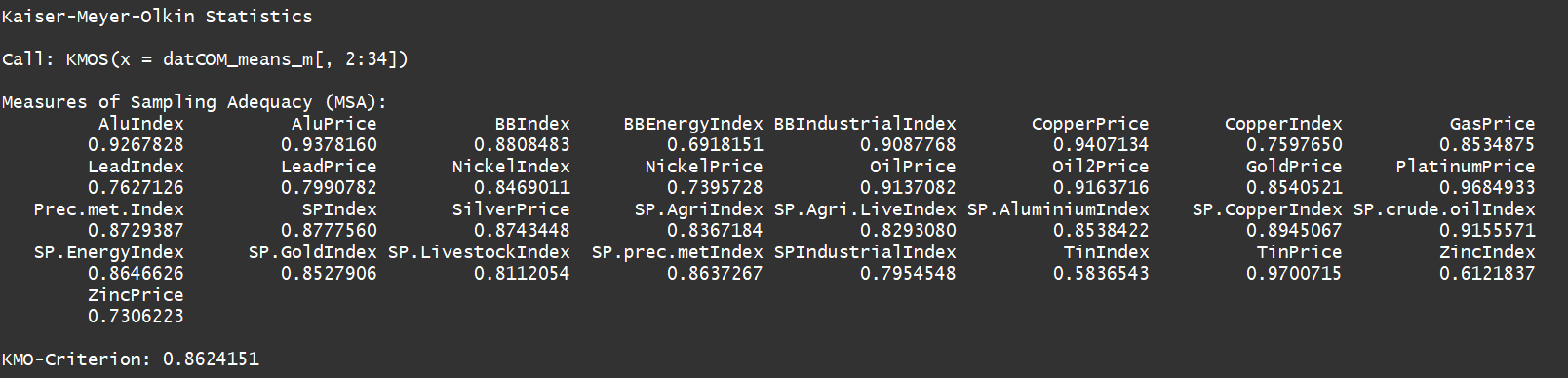
SP.EnergyIndex SP.GoldIndex SP.LivestockIndex SP.prec.metIndex SPIndustrialIndex TinIndex TinPrice ZincIndex

0.8646626 0.8527906 0.8112054 0.8637267 0.7954548 0.5836543 0.9700715 0.6121837

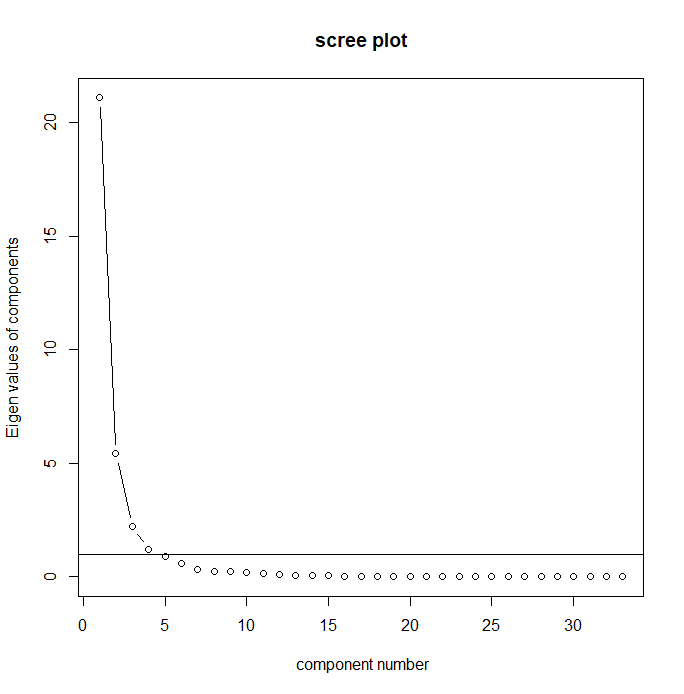
ZincPrice

0.7306223

KMO-Criterion: 0.8624151



We can see that the KMO-Criterion is well above 0.8, therefore the sample size in our data is large enough.



The Screeplot shows us the Eigenvalues of the components that can be used to describe the data. The cut-off point is the eigenvalue of 1. The screeplot indicates that it is likely we need 3 or 4 components. However we will analyse this more concretely in the following steps.

Principal Components Analysis

Call: principal(r = datCOM\_means\_m[, 2:34], nfactors = 10)

Standardized loadings (pattern matrix) based upon correlation matrix

item RC1 RC4 RC2 RC9 RC6 RC3 RC5 RC8 RC10 RC7 h2 u2 com

SP.prec.metIndex 28 0.93 1.00 0.0033 1.3

SilverPrice 19 0.93 0.99 0.0051 1.3

SP.GoldIndex 26 0.93 1.00 0.0044 1.3

Prec.met.Index 17 0.92 0.99 0.0070 1.3

GoldPrice 15 0.92 0.98 0.0220 1.3

SP.Agri.LiveIndex 21 0.90 0.98 0.0180 1.4

TinPrice 31 0.89 0.97 0.0331 1.4

OilPrice 13 0.88 0.99 0.0083 1.6

SP.AgriIndex 20 0.87 0.98 0.0163 1.6

SPIndex 18 0.86 1.00 0.0027 1.7

AluIndex 1 0.85 0.96 0.0402 1.7

Oil2Price 14 0.84 0.99 0.0109 1.9

PlatinumPrice 16 0.84 0.97 0.0253 1.8

SP.crude.oilIndex 24 0.84 0.99 0.0101 1.9

SP.EnergyIndex 25 0.83 0.99 0.0056 1.9

CopperPrice 6 0.80 0.57 0.99 0.0112 1.9

SP.CopperIndex 23 0.80 0.57 0.99 0.0115 1.9

LeadPrice 10 0.78 0.96 0.0438 2.1

SP.LivestockIndex 27 0.78 0.95 0.0541 2.2

SP.AluminiumIndex 22 0.80 0.98 0.0157 2.1

AluPrice 2 0.80 0.99 0.0147 2.1

BBIndustrialIndex 5 0.52 0.79 0.99 0.0086 2.1

NickelPrice 12 0.78 0.97 0.0302 2.3

ZincPrice 33 0.77 0.96 0.0382 2.3

SPIndustrialIndex 29 0.67 0.70 1.00 0.0018 2.2

BBEnergyIndex 4 0.93 0.98 0.0194 1.3

GasPrice 8 0.92 0.95 0.0465 1.3

BBIndex 3 0.71 0.98 0.0188 2.9

ZincIndex 32 0.91 0.98 0.0244 1.4

LeadIndex 9 0.70 0.97 0.0311 3.1

NickelIndex 11 0.77 0.99 0.0148 2.3

TinIndex 30 0.90 0.99 0.0077 1.5

CopperIndex 7 -0.89 1.00 0.0047 1.5

RC1 RC4 RC2 RC9 RC6 RC3 RC5 RC8 RC10 RC7

SS loadings 16.19 6.37 3.81 1.80 1.18 1.11 1.08 0.32 0.27 0.26

Proportion Var 0.49 0.19 0.12 0.05 0.04 0.03 0.03 0.01 0.01 0.01

Cumulative Var 0.49 0.68 0.80 0.85 0.89 0.92 0.96 0.97 0.97 0.98

Proportion Explained 0.50 0.20 0.12 0.06 0.04 0.03 0.03 0.01 0.01 0.01

Cumulative Proportion 0.50 0.70 0.81 0.87 0.91 0.94 0.97 0.98 0.99 1.00

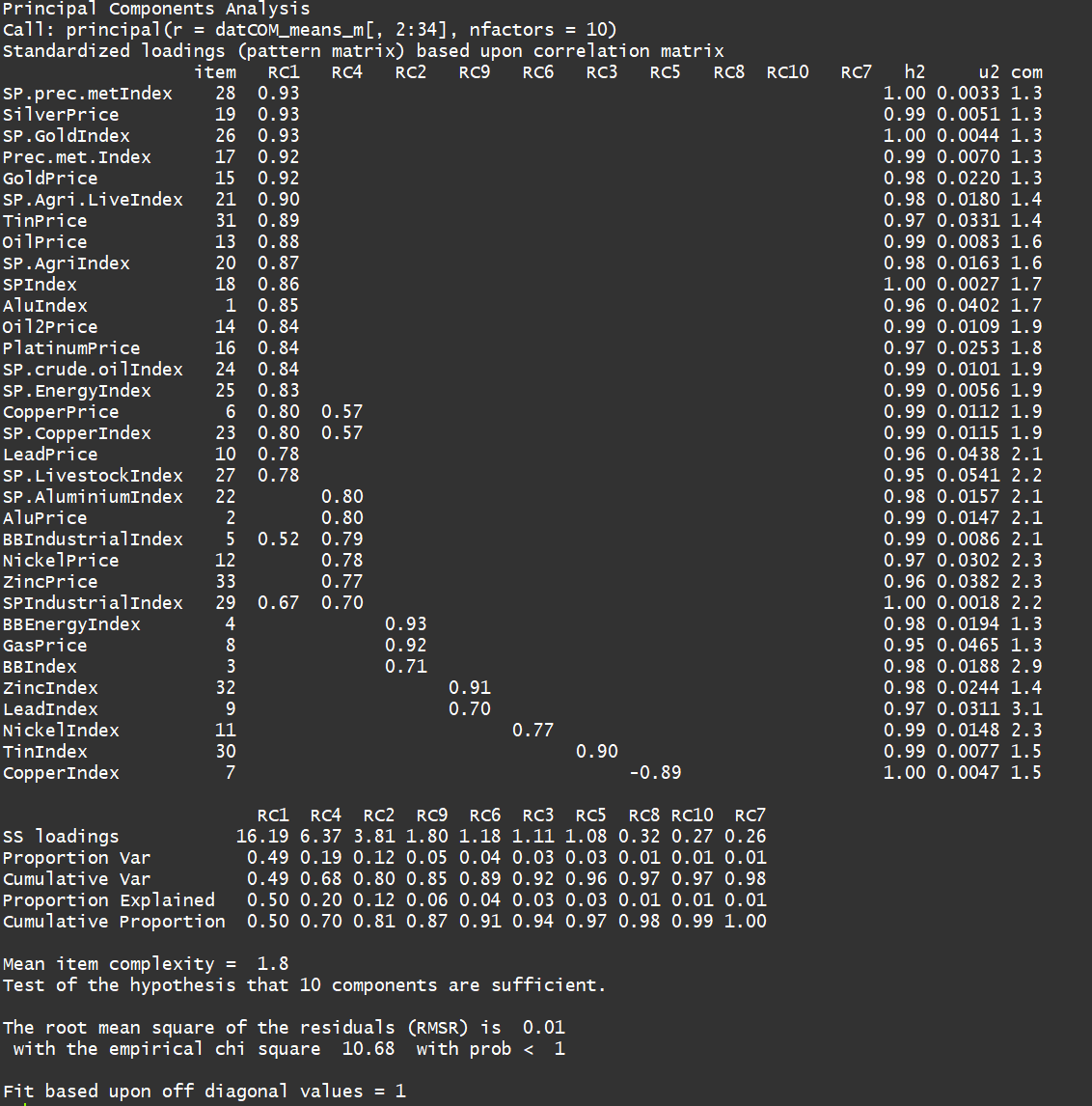
Mean item complexity = 1.8

Test of the hypothesis that 10 components are sufficient.

The root mean square of the residuals (RMSR) is 0.01

with the empirical chi square 10.68 with prob < 1

Fit based upon off diagonal values = 1



We can see in the table above that there are some components that only exist of one variable. Therefore we reduce the amount of components we use to describe the data.

Using 5 components we find

rincipal Components Analysis

Call: principal(r = datCOM\_means\_m[, 2:34], nfactors = 5)

Standardized loadings (pattern matrix) based upon correlation matrix

item RC1 RC4 RC2 RC3 RC5 h2 u2 com

SilverPrice 19 0.94 0.98 0.0239 1.2

SP.GoldIndex 26 0.94 0.97 0.0258 1.2

SP.prec.metIndex 28 0.93 0.97 0.0267 1.2

TinPrice 31 0.90 0.95 0.0451 1.3

Prec.met.Index 17 0.90 0.96 0.0387 1.4

GoldPrice 15 0.89 0.93 0.0675 1.4

SP.Agri.LiveIndex 21 0.87 0.93 0.0677 1.5

SP.LivestockIndex 27 0.87 0.83 0.1666 1.2

OilPrice 13 0.86 0.98 0.0239 1.7

AluIndex 1 0.85 0.89 0.1085 1.5

SPIndex 18 0.84 0.99 0.0080 1.8

Oil2Price 14 0.83 0.97 0.0275 1.9

SP.crude.oilIndex 24 0.82 0.98 0.0230 1.9

PlatinumPrice 16 0.82 0.97 0.0309 1.9

SP.EnergyIndex 25 0.82 0.99 0.0143 2.0

SP.AgriIndex 20 0.81 0.87 0.1281 1.7

LeadPrice 10 0.79 0.53 0.93 0.0743 1.9

CopperPrice 6 0.78 0.60 0.99 0.0127 1.9

SP.CopperIndex 23 0.78 0.60 0.99 0.0130 2.0

NickelIndex 11 0.61 -0.56 0.78 0.2187 2.5

BBIndustrialIndex 5 0.83 0.98 0.0180 1.8

AluPrice 2 0.81 0.95 0.0501 1.9

SP.AluminiumIndex 22 0.81 0.95 0.0520 1.9

NickelPrice 12 0.77 0.88 0.1184 2.0

ZincPrice 33 0.75 0.90 0.0975 2.2

SPIndustrialIndex 29 0.66 0.72 0.99 0.0051 2.2

BBEnergyIndex 4 0.92 0.94 0.0600 1.2

GasPrice 8 0.85 0.87 0.1348 1.4

BBIndex 3 0.51 0.79 0.97 0.0292 2.0

ZincIndex 32 0.87 0.89 0.1149 1.4

LeadIndex 9 0.72 0.86 0.1397 2.3

CopperIndex 7 -0.87 0.91 0.0855 1.4

TinIndex 30 -0.73 0.83 0.1662 2.0

RC1 RC4 RC2 RC3 RC5

SS loadings 15.95 6.92 4.16 2.07 1.68

Proportion Var 0.48 0.21 0.13 0.06 0.05

Cumulative Var 0.48 0.69 0.82 0.88 0.93

Proportion Explained 0.52 0.22 0.13 0.07 0.05

Cumulative Proportion 0.52 0.74 0.88 0.95 1.00

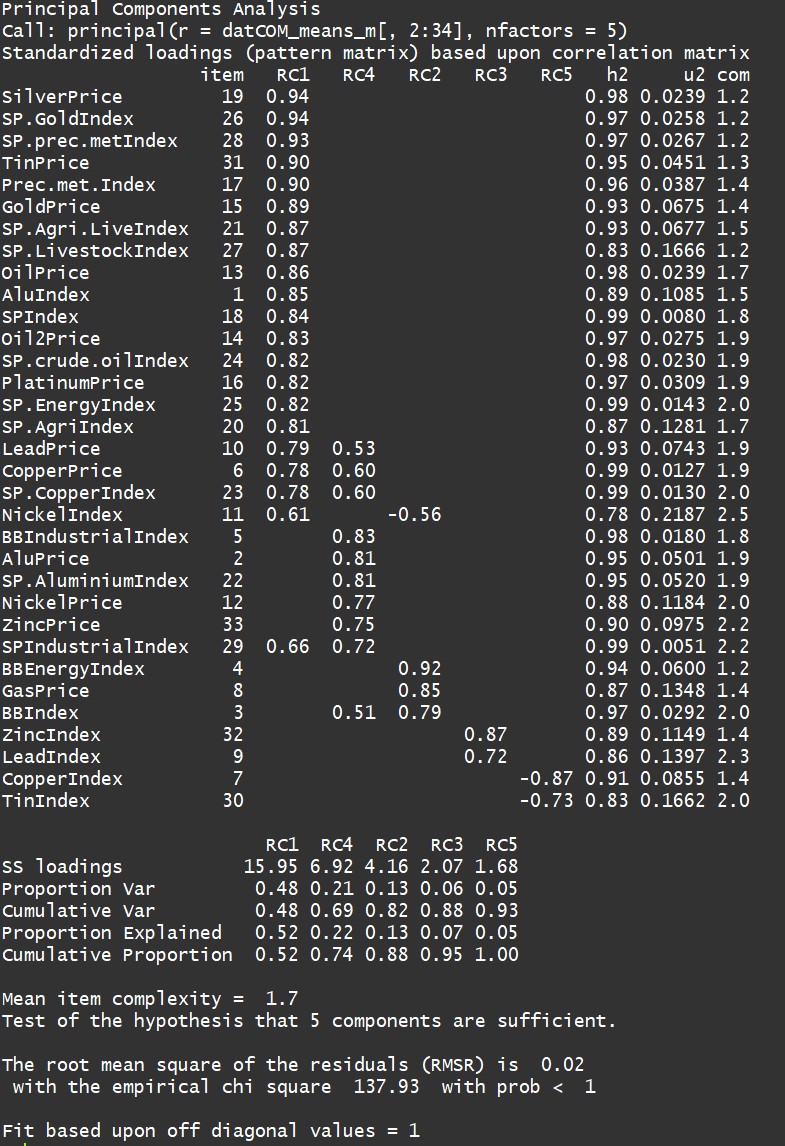
Mean item complexity = 1.7

Test of the hypothesis that 5 components are sufficient.

The root mean square of the residuals (RMSR) is 0.02

with the empirical chi square 137.93 with prob < 1

Fit based upon off diagonal values = 1



We can see that we receive more cross loadings using 5 components. Therefore we reduce the amount of factors/components again.

Principal Components Analysis

Call: principal(r = datCOM\_means\_m[, 2:34], nfactors = 3)

Standardized loadings (pattern matrix) based upon correlation matrix

item RC1 RC2 RC3 h2 u2 com

SilverPrice 19 0.98 0.97 0.030 1.0

SP.prec.metIndex 28 0.98 0.97 0.031 1.0

SP.GoldIndex 26 0.98 0.97 0.031 1.0

Prec.met.Index 17 0.98 0.96 0.041 1.0

TinPrice 31 0.96 0.95 0.050 1.1

GoldPrice 15 0.95 0.93 0.068 1.1

SP.Agri.LiveIndex 21 0.94 0.92 0.078 1.1

SP.AgriIndex 20 0.90 0.86 0.140 1.1

CopperPrice 6 0.90 0.97 0.028 1.4

SP.CopperIndex 23 0.89 0.97 0.028 1.4

LeadPrice 10 0.88 0.91 0.092 1.4

SP.LivestockIndex 27 0.86 0.76 0.239 1.0

OilPrice 13 0.86 0.94 0.056 1.5

SPIndex 18 0.86 0.97 0.031 1.6

AluIndex 1 0.85 0.88 0.122 1.4

PlatinumPrice 16 0.84 0.95 0.049 1.6

SP.crude.oilIndex 24 0.82 0.94 0.060 1.7

Oil2Price 14 0.82 0.93 0.068 1.7

SP.EnergyIndex 25 0.81 0.94 0.056 1.7

SPIndustrialIndex 29 0.80 0.96 0.044 1.9

NickelIndex 11 0.64 0.78 0.224 2.4

ZincPrice 33 0.63 0.78 0.216 2.5

BBIndex 3 0.93 0.94 0.058 1.2

BBEnergyIndex 4 0.86 0.84 0.156 1.3

GasPrice 8 0.85 0.76 0.243 1.1

SP.AluminiumIndex 22 0.71 0.87 0.130 2.2

AluPrice 2 0.71 0.87 0.128 2.2

NickelPrice 12 0.71 0.81 0.188 2.1

BBIndustrialIndex 5 0.63 0.65 0.88 0.118 2.3

TinIndex 30 -0.77 0.64 0.363 1.1

ZincIndex 32 -0.73 0.66 0.338 1.5

LeadIndex 9 -0.63 0.76 0.245 2.6

CopperIndex 7 -0.58 0.44 0.565 1.6

RC1 RC2 RC3

SS loadings 18.66 7.11 2.92

Proportion Var 0.57 0.22 0.09

Cumulative Var 0.57 0.78 0.87

Proportion Explained 0.65 0.25 0.10

Cumulative Proportion 0.65 0.90 1.00

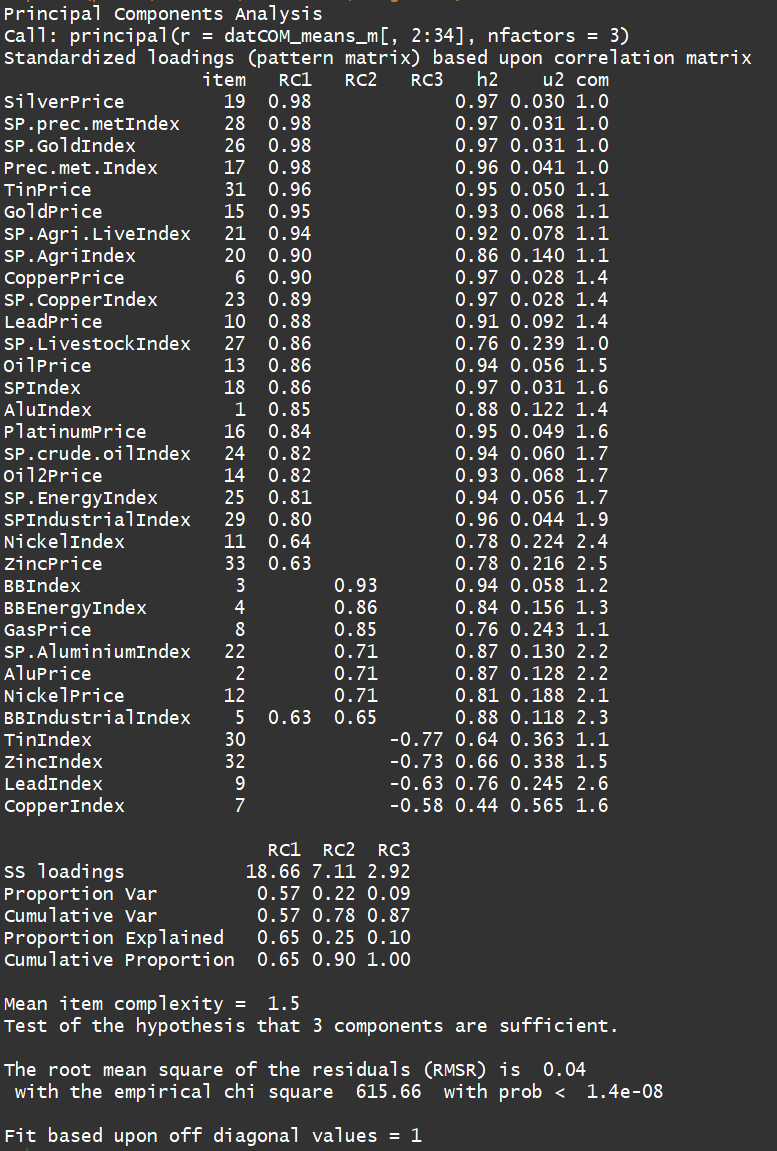
Mean item complexity = 1.5

Test of the hypothesis that 3 components are sufficient.

The root mean square of the residuals (RMSR) is 0.04

with the empirical chi square 615.66 with prob < 1.4e-08

Fit based upon off diagonal values = 1



Eventually we found that using 3 components yielded the best results with the fewest cross-loadings.

After splitting indices and Commodity prices:

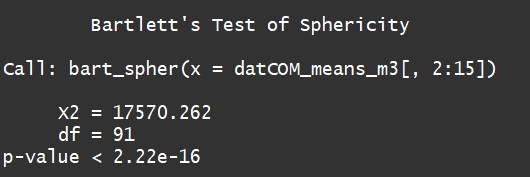
Bartlett's Test of Sphericity

Call: bart\_spher(x = datCOM\_means\_m3[, 2:15])

X2 = 17570.262

df = 91

p-value < 2.22e-16



Kaiser-Meyer-Olkin Statistics

Call: KMOS(x = datCOM\_means\_m3[, 2:15])

Measures of Sampling Adequacy (MSA):

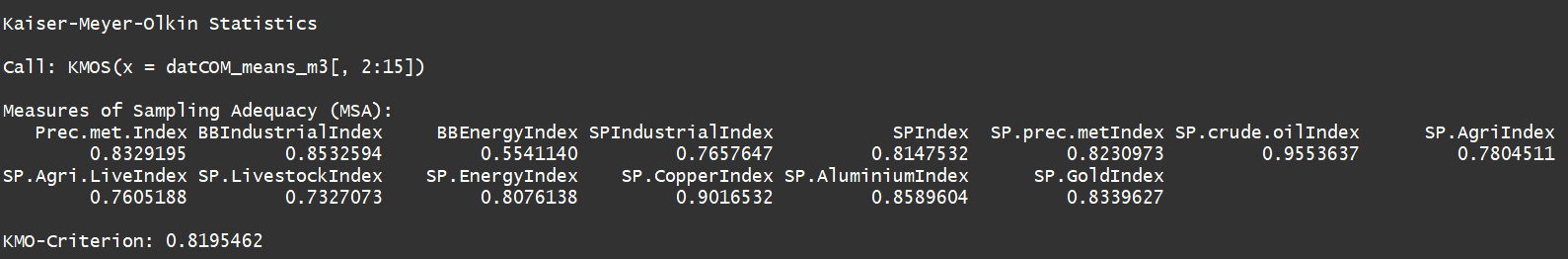
Prec.met.Index BBIndustrialIndex BBEnergyIndex SPIndustrialIndex SPIndex SP.prec.metIndex SP.crude.oilIndex SP.AgriIndex

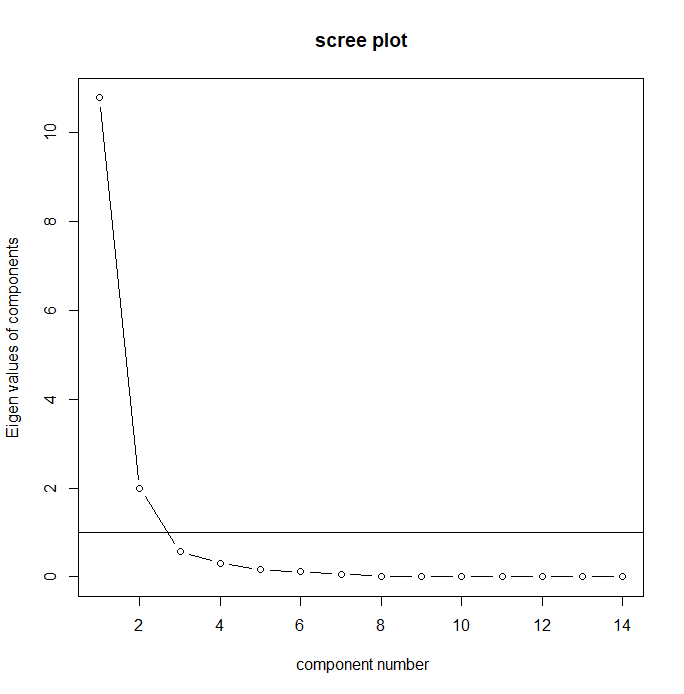
0.8329195 0.8532594 0.5541140 0.7657647 0.8147532 0.8230973 0.9553637 0.7804511

SP.Agri.LiveIndex SP.LivestockIndex SP.EnergyIndex SP.CopperIndex SP.AluminiumIndex SP.GoldIndex

0.7605188 0.7327073 0.8076138 0.9016532 0.8589604 0.8339627

KMO-Criterion: 0.8195462





Principal Components Analysis

Call: principal(r = datCOM\_means\_m3[, 2:15], nfactors = 10)

Standardized loadings (pattern matrix) based upon correlation matrix

item RC1 RC3 RC2 RC4 RC6 RC5 RC7 RC8 RC9 RC10 h2 u2 com

Prec.met.Index 1 0.89 1 7.1e-05 1.5

SP.prec.metIndex 6 0.89 1 2.7e-04 1.6

SP.GoldIndex 14 0.89 1 2.0e-04 1.6

SP.Agri.LiveIndex 9 0.83 1 6.7e-05 2.0

SP.AgriIndex 8 0.82 1 4.9e-05 2.0

SPIndex 5 0.71 1 1.2e-04 2.8

SP.EnergyIndex 11 1 1.9e-04 3.2

SP.crude.oilIndex 7 1 4.4e-04 3.2

SP.AluminiumIndex 13 0.91 1 1.4e-05 1.4

BBIndustrialIndex 2 0.91 1 5.9e-05 1.4

SPIndustrialIndex 4 0.82 1 5.0e-04 1.9

SP.CopperIndex 12 0.70 1 1.4e-04 2.3

BBEnergyIndex 3 0.94 1 2.7e-06 1.3

SP.LivestockIndex 10 0.71 1 3.9e-06 2.2

RC1 RC3 RC2 RC4 RC6 RC5 RC7 RC8 RC9 RC10

SS loadings 6.51 4.52 1.27 0.94 0.42 0.25 0.05 0.02 0.02 0

Proportion Var 0.47 0.32 0.09 0.07 0.03 0.02 0.00 0.00 0.00 0

Cumulative Var 0.47 0.79 0.88 0.95 0.98 0.99 1.00 1.00 1.00 1

Proportion Explained 0.47 0.32 0.09 0.07 0.03 0.02 0.00 0.00 0.00 0

Cumulative Proportion 0.47 0.79 0.88 0.95 0.98 0.99 1.00 1.00 1.00 1

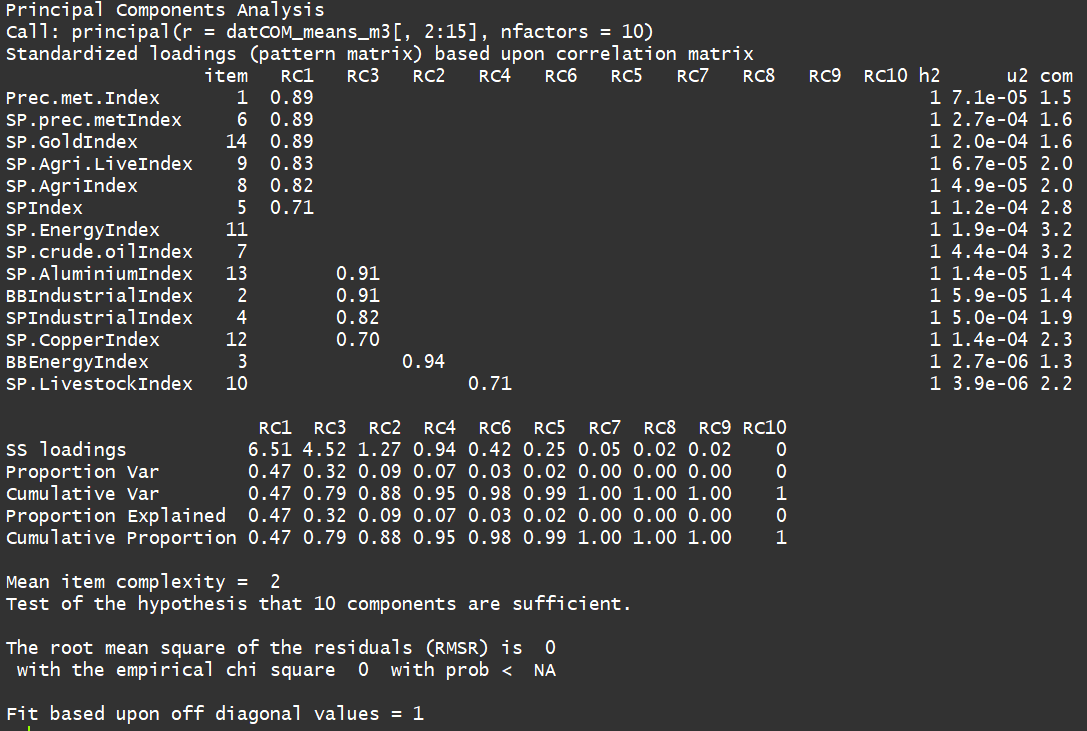
Mean item complexity = 2

Test of the hypothesis that 10 components are sufficient.

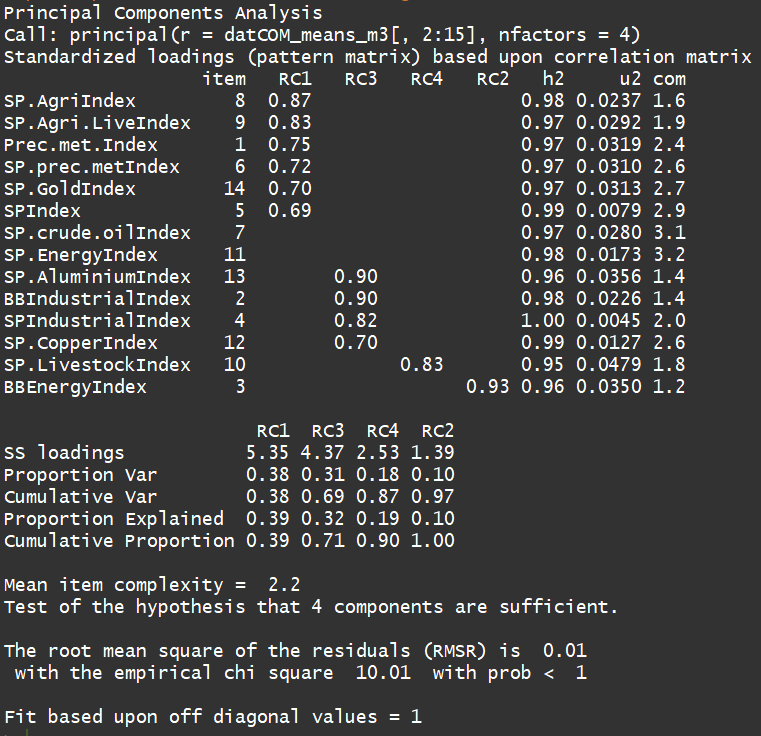
The root mean square of the residuals (RMSR) is 0

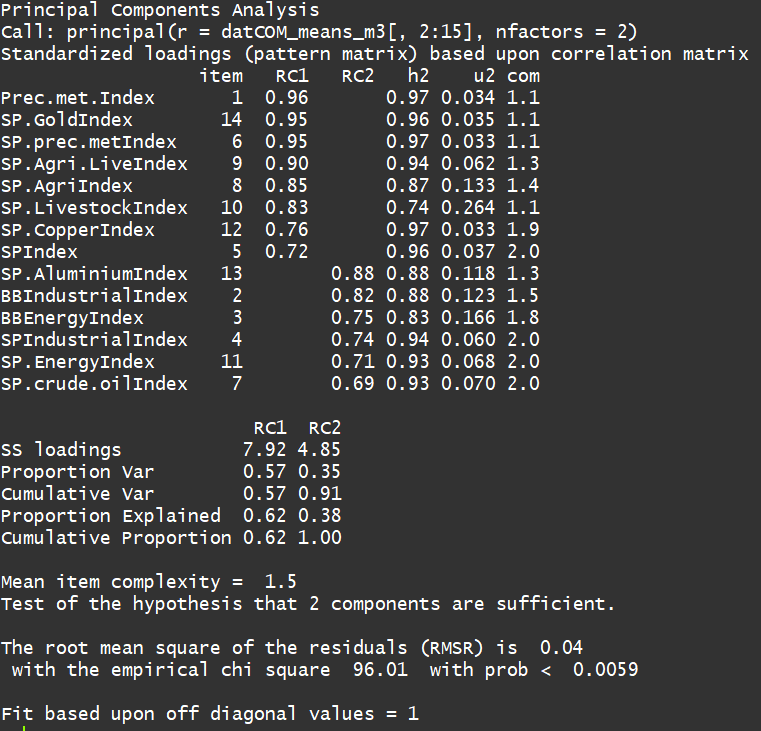
with the empirical chi square 0 with prob < NA

Fit based upon off diagonal values = 1



Reducing the number of Factors:





Eventually we found 2 components, one comprising of precious metals and Agricultural commodities, the other comprising of energy, industry and oil.